Pricing and Hedging Volatility Risk in Fixed Income Markets[*]

Scott Joslin[†]

Current Draft: September 2010

[†]MIT Sloan School of Management. 77 Massachusetts Avenue, Cambridge, MA 02142. `sjoslin@mit.edu`

Abstract

In this paper I show that, with sufficient flexibility in the covariance structure of the risk factors and the market prices of these risks, a low-dimensional term structure model can simultaneously price bonds and related options. I find that a component of volatility risk largely unrelated to the shape of the yield curve is a determinant of expected excess returns for holding long maturity bonds. Moreover, I also find evidence for the converse relationship that the shape of the yield curve affects the premium that agents demand for holding volatility risk. I find that dynamic hedging strategies using bonds alone produce reasonably good hedges for derivative positions during most periods. The structure of risk in my model that gives rise to these features of volatility is distinct from that inherent in recent models with "unspanned stochastic volatility."

1 Introduction

In this paper, I study the joint properties of the risks underlying bond and bond option markets. I find that the volatility of the yield curve is an important predictor of future bond returns and that this volatility is better identified through interest rate options. Moreover, the converse relationship holds as well that the shape of the yield curve identifies the risk premium that investors demand for bearing interest rate volatility risk. I also show that hedging positions sensitive to interest rate volatility risk using bonds requires a dynamic strategy and the degree to which the level, slope, and curvature of the yield curve can hedge such positions is time-varying.

I study these questions through the lens of a four-factor affine term structure model. Though dynamic models with a small number of risk factors (e.g., two or three) have had considerable success at pricing bonds across a broad spectrum of maturities, they typically generate large errors when pricing options on these bonds.[1] There are two critical features of my model that underlie its relative success in simultaneously pricing bonds and bond options. First, I focus on members of the affine family of term structure models ((Duffie and Kan (1996)) that are known to be successful in pricing bonds and allow flexibility in the conditional covariances of the risk factors. In particular, I use an identified version of the affine process specification given in Duffie et al. (2003b) which allows for a richer covariance structure among risk factors than the commonly used specification of Dai and Singleton (2000).[2] The second feature of my analysis is the dependence of the market price of risk on the state of the economy. I follow Cheridito et al. (2007) in parameterizing the market prices of risk, which extends the specification of Duffee (2002) and Dai and Singleton (2002). This extended specification for the market price allows for time-variation in the premium associated with volatility risks, an element that I find critical for matching the data.

I find that the relationship between the yield curve and volatility has a number of interesting features regarding the hedging of portfolio positions

---

[1]Mean-squared relative pricing errors for options on the order of 30% are reported in Buhler et al. (1999), Dreissen et al. (2003), and Jagannathan et al. (2003). Trolle and Schwartz (2009) propose a model with that fits both the term structure of interest rates and a cross-section of options, though their preferred model include a total of 24 factors (18 of which are locally deterministic).

[2]I use the identification scheme in Joslin (2006), but see also Collin-Dufresne et al. (2008).

sensitive to volatility risk. First, I show that the fraction of variation in volatility explained by the the yield curve (as summarized by the level, slope, and curvature) is time-varying where at times over 70% of the variation is explained by variation in the yield while at other times the yield curve explains almost none of the variation times (for example, around the LTCM liquidity crisis). Thus in order to optimally hedge against volatility risk, one must pursue a dynamic strategy whose effectiveness will be time-varying as well. Moreover, in order to hedge interest rate option straddle positions (a position very sensitive to volatility risk), a dynamic strategy is even more critical due to the relative importance of volatility and yield risk as the position moves away from the money.

My model also shows that volatility plays an important role in determining risk premiums that investors demand for bearing interest rate risk. A number of studies have shown that the shape of the yield curve is related to expected excess returns for holding long maturity bonds. My results show that volatility, incrementally to the level, slope, and curvature of the yield curve, is an important determinant of expected excess returns for holding interest rate risk, explaining approximately 40% of the variation in expected returns. This result is consistent with the results of Wright (2009) who argues that inflation uncertainty plays an important role in determining bond risk premiums. This phenomenon offers a potential explanation associated with the 'conundrum' period where during 2004-2007 the federal reserve raised interest rates in 14 straight FOMC meeting while long maturity yields remained relatively constant. As other researcher have noted (e.g. Rudebusch et al. (2006)) this pattern could be attributed to declining risk premiums, one cause of which my model would attribute to declining volatility.

As in equity markets, investors typically demand a risk premium for bearing interest rate volatility risk. This premium is typically negative since volatility is usually high in bad states, though the premium does occasionally become positive. I demonstrate a two-way feedback effect where, in addition to the effect volatility has on the premium for bearing interest rate risk, there is an effect where the shape of the yield curve helps determine the premium that investors demand for bearing volatility risk.

Although my model incorporates a component of volatility risk that varies independently of the level, slope, and curvature of the yield curve, the mechanism is very different from a model with unspanned stochastic volatility (USV, see Collin-Dufresne and Goldstein (2002b)). In these models, volatility varies independently of the entire yield curve due to a very specific type

4

of cancellation. In general, volatility will drive long maturity interest rates through two channels: (i) a convexity effect and (ii) through an expectations effect whereby changes in the level of volatility affects (risk-neutral) expectations of future short rates. Models with the USV property rely on an exact cancellation of these two effects across maturities. I show that the first channel, in fact, generates very little variation in the yield curve because convexity effects are very small for short maturities while mean reversion of volatility implies that the convexity effect at long maturities is nearly constant. In my estimation, I consider the most general model in order to let the data select the necessary ingredients in the model. I find a component of volatility which, while generating small variations in convexity effects across maturities, also has very little effect on risk neutral expectations of future short rates. Under these conditions, a component of volatility will very little effect on the shape of the yield curve. As I elaborate further in Section 7, such a model turns out to be quite different from a model where volatility affects expectations in such a way to exactly cancel (across all maturities) the convexity effects that it generates.

From a methodological perspective, essential to exploring the issues addressed in this paper is an ability to compute the prices of options, for which closed-form solutions do not exist, and the joint conditional likelihood function of a large cross-section of bond yields and option prices. I develop a Fourier analytic quadrature technique for computing option prices. I also extend this technique to develop a feasible method for full information maximum likelihood estimation of affine diffusions. These results are applicable to a wide variety of problems beyond those examined in this paper, both in bond and equity markets, and therefore they are potentially of interest in their own right.

The remainder of the paper is organized as follows. Section 2 describes the model and estimation procedure. Section 3 provides a summary of the estimation results. The hedging of volatility risk is discussed in Section 4. The pricing of yield and volatility risk are examined in Section 5 and Section 6. Section 7 considers the role of convexity in bond prices. Finally, Section 8 concludes.

2 Model

I consider 4-factor affine short-rate models.[3] The short rate, r_t, is driven by a state variable, X_t, such that

$$r_t = \rho_0 + \rho_1 \cdot X_t, \qquad (1)$$

and

$$dX_t = \mu_t^{\mathbb{P}} dt + \sigma_t dB_t^{\mathbb{P}},$$
$$\mu_t^{\mathbb{P}} = K_0^{\mathbb{P}} + K_1^{\mathbb{P}} X_t, \qquad (2)$$

where $\rho_1, K_0^{\mathbb{P}} \in \mathbb{R}^4$, $K_1^{\mathbb{P}} \in \mathbb{R}^{4\times4}$, and $B_t^{\mathbb{P}}$ is a standard 4-dimesional Brownian under $\mathbb{P}$, the historical measure. Duffie et al. (2003b) give conditions for (2) to give a well-defined process on $\mathbb{R}_+^M \times \mathbb{R}^{N-M}$. Here the covariance is given by $\sigma_t \sigma_t' = \Sigma_0 + \sum_{i=1}^M \Sigma_i X_t^i$. I consider $A_M(4)$ models where either $M = 1$ or $M = 2$ factors drive volatility. For example, in the $A_2(4)$ case this means that, $\sigma_t \sigma_t' = \Sigma_0 + \Sigma_1 X_t^1 + \Sigma_2 X_t^2$, a 4×4 matrix. The constraints in Duffie et al. (2003b) require that (i) each Σ_i is positive semi-definite, (ii) $\Sigma_{1,22} = \Sigma_{2,11} = 0$, (iii) $K_{1,ij}^{\mathbb{P}} = 0$ for $i \le 2$ and $j \ge 2$, (iv) $K_{1,12}^{\mathbb{P}}, K_{1,21}^{\mathbb{P}} > 0$ and (v) $K_{0,1}^{\mathbb{P}}, K_{0,2}^{\mathbb{P}} > 0$. These conditions insure that the covariance is always positive semi-definite and the first two factors, which drive volatility, always remain positive. As Joslin (2006) notes, in the $A_2(4)$ case, this specification allows for greater flexibility in the correlation structure among the risk factors than the normalization of Dai and Singleton (2000).

The dynamics of the economy are linked to the pricing measure by the market prices of risk. I use the completely affine market price of risk specification in Cheridito et al. (2007). This specification allows the expected excess returns for exposure to each risk factor to be affine in the state. As elaborated further in Section 7, a flexible market price of risk is critical in matching observed risk premia for holding both bonds and bond options. Under this market price of risk specification, the dynamics of the state variable X_t are affine under $\mathbb{Q}$ as well and satisfy

$$dX_t = \mu_t^{\mathbb{Q}} dt + \sigma_t dB_t^{\mathbb{Q}},$$
$$\mu_t^{\mathbb{Q}} = K_0 + K_1 X_t, \qquad (3)$$

---

[3]See Dai and Singleton (2000) for a summary of affine term term structure models. They classify affine term structure models into non-nested families denoted $A_M(N)$. N is the total number of factors and M is the number of factors driving volatility.

where $B_t^{\mathbb{Q}}$ is a 4-dimensional standard Brownian motion under $\mathbb{Q}$ and K_0 and K_1 satisfy the same conditions as before. The absence of arbitrage is then guaranteed by assuming that the Feller condition is satisfied under both measure so that $K_{0,i}^{\mathbb{P}} \geq \frac{1}{2}\Sigma_{i,ii}$ and $K_{0,i}^{\mathbb{Q}} \geq \frac{1}{2}\Sigma_{i,ii}$ for $i \leq M$.

In order to ensure the parameters are econometrically identified, I impose the normalization constraints given in Joslin (2006). See Appendix A for further details.

Any claim with payoff at time T given by $f(X_T)$ can be priced by the discounted risk-neutral expected value

$$E_t^{\mathbb{Q}}\left[e^{-\int_t^T r_\tau \mathrm{d}\tau} f(X_T)\right]. \tag{4}$$

Duffie and Kan (1996) show that zero coupon bond prices are given by

$$P_t^T(X_t) = e^{A(T-t)+B(T-t)\cdot X_t}, \tag{5}$$

where P_t^T denotes the price at time t for a zero coupon bond paying \$1 at time T. The loadings A and B satisfy the Riccati differential equations

$$\dot{B} = -\rho_1 + \left(K_1^{\mathbb{Q}}\right)' B + \frac{1}{2}B^\top H_1 B, \quad B(0) = 0, \tag{6}$$

$$\dot{A} = -\rho_0 + \left(K_0^{\mathbb{Q}}\right)' B + +\frac{1}{2}B^\top \Sigma_0 B, \quad A(0) = 0, \tag{7}$$

where H_1 is a tensor in $\mathbb{R}^{4\times4\times4}$ defined (as in Duffie (2001)) so that $B'H_1B$ is a 4-dimensional vector with $(B'H_1B)_i = B'\Sigma_i B$.

Collin-Dufresne and Goldstein (2002b) show that it is possible that some linear combination of the bond loadings is identically zero for all maturities. In such case, a volatility factor can affect conditional second moments but not be contemporaneously spanned by bonds. Such unspanned volatility factors will directly affect fixed income derivative prices. I therefore estimate models with the additional constraints required for unspanned stochastic volatility in in addition to the more general specifications.

I also consider both interest rate caps and swaptions. An interest rate cap is a portfolio of options on 3-month LIBOR that caps the interest rate paid on a floating loan. An interest rate swap is an option to enter into a swap, exchanging a fixed interest rate for a floating interest rate. Since the floating side of the swap is always worth par, a swaption is equivalent to an option on a coupon bond.

An option on a Q-year swap expiring in P-year, referred to as an in-P-for-Q swaption, may be priced by

$$S_t = E^{\mathbb{Q}}[e^{-\int_t^{t+P} r_\tau d\tau}(CB(X_{t+P}, Q) - 1)^+] \,, \tag{8}$$

where $CB(X, Q)$ is the price when the state is X of a Q-year coupon bond with coupon equal to the strike. Singleton and Umantsev (2003) approximate this expectation replacing the exact exercise region, $\{CB(X_{t+P}, Q) \geq 1\}$, with the region implied by a linearization of the swap rate. Since the coupon bond price is a sum of coupons whose prices are exponential affine functions of the state, this reduces the problem of pricing the swaption to that of computing forward probabilities which may be evaluated by the transform method in Duffie et al. (2000).[4] An interest rate caplet then becomes a special case where the linearization is exact.

In estimation of the models, computation is required for a large number of caps and swaption coupons. This involves evaluations of many transforms each of which is an integral whose integrand is defined as the solution of an ODEs similar to (6–7) which must be solved numerically for the general models that I consider. Because of this difficulty, I develop an adaptive integration scheme to compute the required forward probabilities. This scheme gives very accurate prices using only 3 or 4 quadrature nodes. See Appendix B for details.

After computing pricing securities using the dynamics under the risk neutral measure, it remains to estimate the parameters governing the evolution of the economy under the physical measure. Ideally, one would like to estimate the affine diffusion in equation (2) by maximum likelihood. Although the exact transition likelihood for an affine diffusion is known in terms of Green's functions of the Feynman-Kac PDE, direct computation is intractable. There is a very extensive literature which deals with alternative estimations methods. Some alternative approaches to maximum likelihood include moment-based estimators (e.g. QML, GMM, or characteristic-function based methods as in Singleton (2001), Carrasco et al. (2006), and others), simulation methods (e.g. Duffie and Singleton (1993) and Brandt and Santa-Clara (2002)), and

---

[4] Collin-Dufresne and Goldstein (2002a) suggest computing swaption prices using an Edgeworth expansion using the cumulants of the price of the associated coupon bond. This approach presents a potential problem that Edgeworth expansions do not in general converge. Additionally, to compute the k-th moment of a 10-year coupon bond with semi-annual coupon requires the numerical solution of $\binom{20+k-1}{k}$ differential equations. For $k = 6$ this already reaches 177,100 equations.

approximate methods (e.g. Duffie et al. (2003a), Ait-Sahalia (1999), and Ait-Sahalia and Kimmel (2010)). However, I estimate the models using full information maximum likelihood estimation using an extension of the methods that I develop for pricing options. See Appendix C for a summary of the calculations used in the current context.

I compute the likelihood of the observed time series of data as follows. First, for each panel observation of zero coupon yields and option prices, I suppose that three zero coupon yields and one swaption price are observed exactly. Given these prices and the underlying parameters, I can then invert the state using Newton's method.[5] The likelihood of the prices of these instruments assumed to be priced exactly then is computed by computing the likelihood of the inverted state (as in Appendix C) and applying the Jacobian of the linearized transformation at the observed state. The likelihood of the complete data is then computed with the assumption that all other yields and option prices are assumed to be observed with i.i.d. normal measurement errors.

3   Estimation Results

The data, obtained from Datastream, consists of LIBOR, swap rates, and at-the-money swaption and cap implied volatilities from June 1997 to June 2006. I use 3-month LIBOR and the entire term structure of swap rates to bootstrap swap zero rates. The bootstrap procedure assumes that forward swap zero rates are constant between observations.

The models are estimated using 6 month, 1-,2-,3-,4-,5-,7-, and 10-year swap-zero rates. The models are also estimated using swaptions with expiries of 3 months, 1 year and 3 years written on swaps with maturities of 2 years, 5 years, and 8 years. Interest rate caps with maturities of 2 years, 5 years, and 8 years are also used in estimation. The models assume that the 6-month, 2-year, and 10-year yields are priced without error along with the 1 year into 5 year swaption. The remaining instruments are priced with errors which are assumed to be independent and normally distributed.

The model estimates are given in Table 1 and Table 2. Throughout, the superscript USV in the model name refers to the estimated model where the unspanned stochastic volatility constraints are imposed.

---

[5]The pricing relation was more nearly linear to equate the model implied Black volatility.

	$A_1(4)$	$A_2(4)$	$A_1(4)^{USV}$	$A_2(4)^{USV}$
$K_{0,1}^{\mathbb{Q}}$	0.5247	0.5	0.5	1.751
$K_{0,2}^{\mathbb{Q}}$	0	0.5	0	1.054
$K_{1,11}^{\mathbb{Q}}$	-0.3174	-0.8255	-0.2439	-3.833
$K_{1,12}^{\mathbb{Q}}$	0	0.4036	0	0
$K_{1,21}^{\mathbb{Q}}$	0.08395	0.8716	0.1031	0.648
$K_{1,22}^{\mathbb{Q}}$	-0.6417	-1.085	-0.03375	-1.628
$K_{1,23}^{\mathbb{Q}}$	0	0	-0.6808	0
$K_{1,24}^{\mathbb{Q}}$	0	0	-3.998	0
$K_{1,31}^{\mathbb{Q}}$	0.3874	1.673	-0.02381	-1.217
$K_{1,32}^{\mathbb{Q}}$	0	-1.491	0	0.08898
$K_{1,33}^{\mathbb{Q}}$	-0.05028	-0.6241	-0.0675	-0.1183
$K_{1,34}^{\mathbb{Q}}$	0	0	-0.9504	-0.4021
$K_{1,41}^{\mathbb{Q}}$	0.1344	0.7582	0	-1.069
$K_{1,42}^{\mathbb{Q}}$	0	0.03494	0	-0.03562
$K_{1,43}^{\mathbb{Q}}$	0	0	0	0
$K_{1,44}^{\mathbb{Q}}$	-1.492	-0.05267	-1.233	-0.2366

Table 1: Drift Parameter Estimates

This table provides model estimates drift parameter estimates. All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed. Unreported parameters are set to zero by the normalization constraints.

	$A_1(4)$	$A_2(4)$	$A_1(4)^{USV}$	$A_2(4)^{USV}$
$\Sigma_{0,22}$	8.951	0	1	0
$\Sigma_{0,32}$	5.518	0	0	0
$\Sigma_{0,33}$	3.411	0.1914	1	0.01838
$\Sigma_{0,42}$	3.699	0	0	0
$\Sigma_{0,43}$	2.281	-0.4824	0	0.1468
$\Sigma_{0,44}$	1.534	1.217	1	2.136
$\Sigma_{1,11}$	1	1	1	1
$\Sigma_{1,22}$	4.426	0	6.925	0
$\Sigma_{1,32}$	1.35	0	0	0
$\Sigma_{1,33}$	8.553	0.6444	0	0.02651
$\Sigma_{1,42}$	2.473	0	0	0
$\Sigma_{1,43}$	-1.401	-2.763	0	0.02399
$\Sigma_{1,44}$	2.029	11.87	0	5.057
$\Sigma_{2,22}$	0	1	0	1
$\Sigma_{2,33}$	0	1.415	0	2.153
$\Sigma_{2,43}$	0	-4.504	0	0
$\Sigma_{2,44}$	0	16.14	0	0

Table 2: Variance Parameter Estimates

This table provides model estimates variance parameter estimates. All matrix are symmetric, with the lower diagonal reported. The $A_1(4)^{USV}$ model has scale normalization through H_0 rather than ρ_1. All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed. Unreported parameters are set to zero by the normalization constraints.

Table 3 presents the root mean square pricing errors for zeros. For the maturities included in the estimation, pricing errors range from 5-10 basis points with the USV models having slightly higher pricing errors. Also tabulated are pricing errors for maturities over 10 years, which were not used in estimation. We discuss these results further in Section 7

Table 4 provides the pricing errors for swaptions and caps. Data from GovPX indicates that swaption bid-ask spreads range from 1-2% implied volatility and about 1%. The slightly higher pricing errors in the short maturity-short expiry options occur mainly during periods of very low interest rates. For example, if the period when the 6 month rate is less than 2%, the mean square error on the in 3 months-for 2 year swaption drops to 2.8%. Thus there is in general a good cross sectional fit across the options as well as the yields.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
1 Year	7.4	12.7	7.4	10.3
2 Year	0.0	0.0	0.0	0.0
3 Year	4.1	10.3	4.1	6.3
4 Year	5.2	15.1	5.2	8.2
5 Year	5.3	16.4	5.3	8.3
7 Year	3.8	13.0	3.8	6.1
10 Year	0.0	0.0	0.0	0.0
12 Year	3.9	12.9	3.9	6.9
15 Year	8.9	38.8	9.1	19.7
20 Year	12.9	96.4	13.2	39.0
25 Year	13.3	176.0	13.6	53.0
30 Year	17.3	279.1	17.4	66.2

Table 3: Zero Coupon Pricing Errors
Root mean square zero coupon yield pricing errors in basis points. Zero coupon yields are computed by bootstrapping the swap curve. All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
3 Months into 2 Years	4.1	7.2	4.1	21.5
3 Months into 5 Years	1.5	2.1	1.4	5.4
3 Months into 8 Years	1.4	1.5	1.4	3.1
1 Year into 2 Years	1.0	3.6	0.9	5.7
1 Year into 5 Years	0.0	0.0	0.0	0.0
1 Year into 8 Years	0.5	0.5	0.5	0.6
3 Years into 2 Years	1.0	1.5	0.8	2.0
3 Years into 5 Years	0.7	0.8	0.7	1.8
3 Years into 8 Years	0.9	0.8	0.8	1.9

Table 4: Swaption Implied Volatility Errors

Root mean square errors in swaption implied volatility errors. Swaptions are considered to be at-the-money in the model. All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed.

To understand the role of risk premia in matching both markets, observe that the likelihood is made up of a component due to the transition dynamics of the economy and a component due to pricing errors. The pricing component is determined by the risk neutral drift ($\mu^{\mathbb{Q}}$) and covariance structure (σ) of the risk factors, while the likelihood of the data measured without error is determined by the drift under the physical measure($\mu^{\mathbb{P}}$) and the covariance structure. The drift under the two measures is related by the market prices of risk. Thus the covariance structure provides a link between the two the likelihood of the pricing errors and the likelihood of the data measured without error.

As elaborated in Section 7, convexity only a small role in bond prices. This means that bond prices depend primarily on risk neutral expectation, $\mu^{\mathbb{Q}}$. Provided the market price of risk is not restrictive, the likelihood cannot be dominated by the pricing errors and the model will be estimated in a consistent manner. On the other hand, with a constrained market price of risk, there will be a tension between the dynamics and pricing errors (related points are discussed in Dai and Singleton (2003)). The completely affine market price of risk allows for risk premia to depend on the state in two important ways. First, it allows for risk premia to depend on the slope of the yield curve and change sign over time. Second, it also allows the risk premium demanded for holding volatility risk to not shrink to zero as volatility drops to zero – that is, investors may still be averse to volatility risk, even when volatility is low.

The risk premium for volatility risk is particularly important in matching the cross-section of option prices. Agents are exposed to interest rate risks directly through holding bonds and also indirectly through asset prices linked to interest rates, such as home values. When interest rate volatility is low, these assets become less risky. This means that when volatility is low, an increase in volatility turns a portion of the investor's portfolio from a riskless asset to a risky asset. If the price of volatility risk is proportional to the level of volatility, the agent is effectively close to risk neutral to changes in the risk-level of large portions of their portfolio.

4   Hedging Volatility Risk

Within the models considered, only the first (for the $A_1(4)$ models) or the first two factors (for the $A_2(4)$ models) affect the volatility of the yield curve.

15

On the other hand, these volatility factors will in general also affect the shape of the yield curve through any of the three terms in (6). That is, the factors driving volatility may jointly drive interest rates by (i) directly entering into the short rate through non-zero ρ_1 loadings in (1), (ii) affecting risk-neutral expectations of the risk-factors through non-zero entries in the corresponding rows of $K_1^{\mathbb{Q}}$, or (iii) through stochastic convexity effects in Σ_i.

To study the effects of volatility risk that are largely unrelated to the yield curve, I consider the component of variance risk which is locally uncorrelated with the level, slope, and curvature of the yield curve. For this purpose, define the residual variance

$$V_t^R = V_t - \alpha_1 \ell_t - \alpha_2 s_t - \alpha_3 c_t, \tag{9}$$

where V_t is the variance of the 5-year zero and (ℓ_t, s_t, c_t) are measures of the level, slope and curvature of the yield curve: $\ell_t = \frac{1}{3}(y_{6m} + y_{2y} + y_{10y})$, $s_t = y_{10y} - y_{6m}$, $c_t = -y_{6m} + 2y_{2y} - y_{10y}$. α is chosen so that V_t^R is locally uncorrelated with (ℓ_t, s_t, c_t): $\alpha = \Sigma_y^{-1} \Sigma_{Vy}$.[6] Here, it is convenient to work with variance instead of volatility since variance is an affine function of the state. There will be two time scales which will be relevant for the residual variance. For example, if one purchases an in 1-year for 5-year swaption with the intention of selling it in three months, they will be concerned with the 9-month volatility of the swap rate 3 months from the purchase. The local residual variance refers to annualized limit when both time scales go to zero. Since the covariance of the factors is time-varying, the weights α are time-varying as well. In the case of a general affine model without USV imposed, the residual variance will have a direct effect on the yields. In such models, the yield curve identifies the volatility exactly. In this sense, the residual volatility incorporates both volatility and the residual risk in the level of interest rates themselves.

Implicit in this definition is the idea that the residual variance primarily drives volatility of the yield curve rather than the shape of the yield curve. In the case of the models with USV, the residual variance factor in fact has exactly no incremental effect on the yield curve. This turns out to be approximately true also in the case for the unconstrained model. Figure 1 plots the effect of changes in the local residual variance, fixing the 6-month, 2-year, and 10-year yields, on the cross section of yields for the estimated $A_1(4)$

---

[6] Alternatively, one could use different maturities or use principal components (as in Joslin et al. (2010b)) with the volatility of the level factor instead of individual maturities.

16

model.[7] The effect of the residual variance on the yield curve is non-zero, but quite small with a one standard deviation weekly shock resulting in a shift of less than half a basis point in all but very short maturity yields. This indicates again that, although the model does not precisely have unspanned volatility, the residual variance, and thus volatility itself, is only very poorly identified from the cross section of bond yields. These results agree well with Litterman and Scheinkman (1991), who show that three principal components explain nearly all of the variation in the yield curve. Thus one would anticipate that a fourth factor likely would have only a small effect on yields. Duffee (2010) finds support, within the context of a Gaussian model, for a factor which has a small effect only on returns. In my model, as I will show, this fourth factor drives both volatility *and* expected excess returns. I elaborate further on the mechanism that generates this effect in comparison to the restrictions in the USV model in Section 7.

Figure 2 plots the time series of local correlation of residual variance with variance of the 5-year yield for the $A_1(4)$ model. A high correlation between the variance and residual variance indicates that the yield curve, as summarized by its level, slope and curvature, is explaining little of the variation in the yield volatility. The correlation typically ranges from 40% to 60% indicating that about half of the variation in the variance of the 5-year yield is accounted for by factors unrelated to the level slope and curvature of the yield curve. A notable exception it that in late 1998, around the time of Russian default and LTCM bailout, the model indicates that the variance was nearly perfectly correlated with the residual variance. This indicates that in this period movements in the volatility of the yield curve were almost completely uncorrelated to movements in the level, slope, and curvature of the yield curve.

To understand the impact in terms of asset prices of the relationship between residual variance and the yield curve, we consider the effect of residual variance on swaption prices. Figure 3 plots the fraction of variation in the in 1 year-for 5 year swaption price due to residual variance risk. For the swaption itself, the residual variance accounts for almost none of the variation in the swaption price. In other words, the delta risk (exposure to changes in the underlying) in a single swaption is much larger than vega risk (exposure to

---

[7]These loadings are found by transforming the original risk factors X_t from the drift-normalize model to the risk factors $Y_t = (\ell_t, s_t, c_t, V_t^R) = C + DX_t$. The new loadings for maturity τ are transformed by $B(\tau) \mapsto (D^{-1})'B(\tau)$

changes in volatility). However, also plotted is the fraction of variance in the quoted prices of the refreshed at-the-money swaption. Here, the moneyness of the swaption is changed to reflect the new at-the-money strike as the yield curve moves through time. This does not reflect the price process available to an investor (since when an investor buys a swaption the strike is fixed), but instead represent the time series of implied volatility translated into a price and allows us to isolate the effect on price due to changing volatility by removing the effect on price of changing moneyness. When the effect of changing moneyness is removed, the residual variance explains a much larger portion of the variation typically from 20%-40%, but as high as 70%. This fraction is nearly the same as with the fraction of variance in straddle prices explained by the residual variance. These results suggest that, at least locally, a hedge of a swaption straddle using bonds will be moderately successful.

Figure 4 shows the sensitivity of a straddle position which is long both a put and a call on the 5-year coupon bond with one year to expiration. The sensitivity is plotted as the moneyness of the options is varied from 30 basis points out of the money to 30 basis points in the money. Typical weekly volatility for the 5-year swap rates range from 12 to 18 basis points, so the range plotted would coincide with approximately a two standard deviation movement (up or down) over a one week period. Over a one month period, the graph denotes approximately a one standard deviation movement, so over this period a movement in the moneyness moving outside the range will occur fairly often, approximately 30% of the time. As the straddle goes away from the money, it becomes much less sensitive to volatility risk. This stands in contrast to equity options, where the volatility of volatility is much larger relative to the volatility of the underlying. From January 2000 to August 2006, S&P 500 index ranged from 776 to 1509 with VIX ranging from 10.2% to 42.1%. The weekly standard deviation of SPX and VIX were 168 points and 6.86%, respectively. Thus in the case of SPX options, volatility risk is a much more important component.

This analysis suggests that a dynamic hedging strategy is particularly important in hedging a swaption straddle. When initiated at the money, the straddle is exposed primarily to volatility risk which may be partially hedged using bonds. After the straddles moves away from the money, the position becomes much more sensitive to the underlying yields relative to the volatility risk. It is important to note also that bid-ask spreads are typically relatively high for fixed income derivatives. This suggests that hedging over a very short horizon will very likely need to rely on the partial hedge available through

18

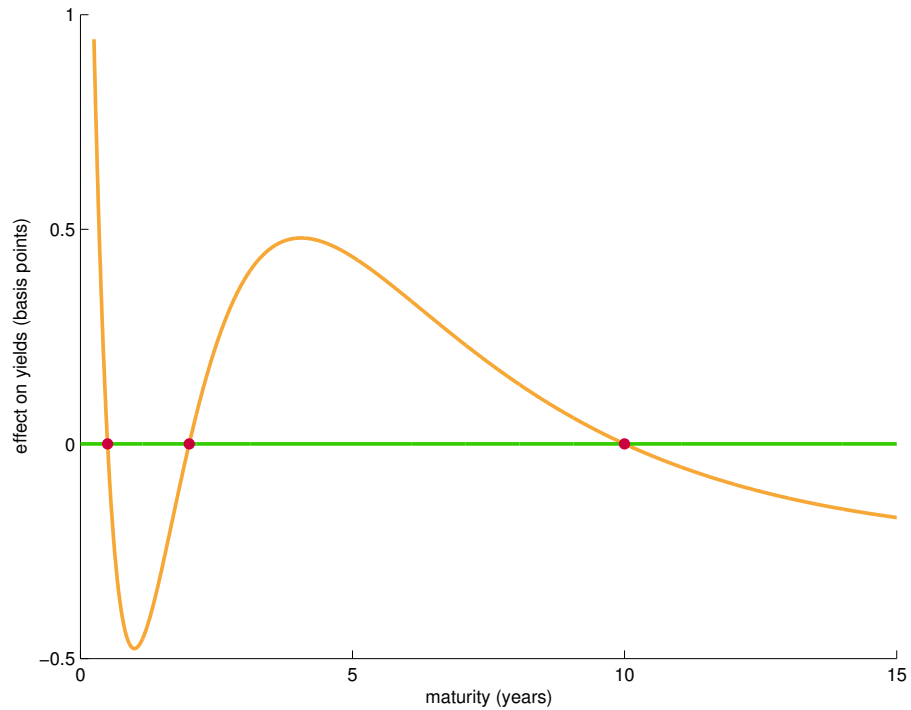the correlation with swaps which have extremely low trading costs.

Figure 1: Effect of Residual Variance on the Yield Curve
This figure plots the effect of the residual variance on the yield curve. The residual variance is defined as the risk which is locally uncorrelated with the 6-month, 2-year, and 10-year yields. The figure plots the effect of a weekly one standard deviation shock in the residual variance on the yield curve, fixed the 6-month, 2-year, and 10-year yields for the $A_1(4)$ model on June 21, 2006.
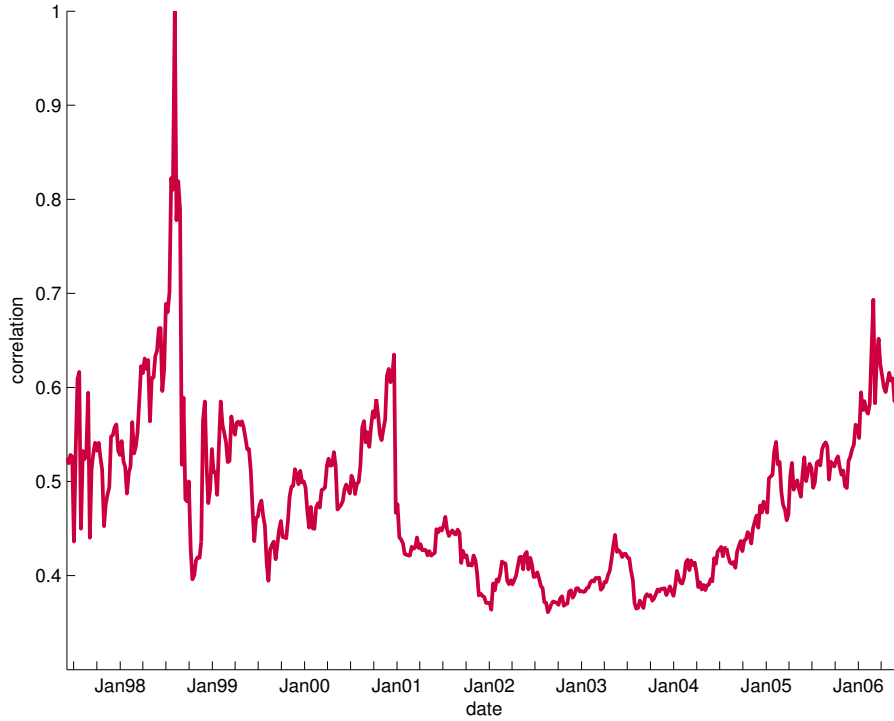
Figure 2: Correlation of Variance with Residual Variance

This figure plots the correlation of the variance of the 5-year yield with the residual variance for the $A_1(4)$ model. The residual variance is defined as the risk which is locally uncorrelated with the 6-month, 2-year, and 10-year yields. A correlation of 1 between the variance and residual variance indicates that 6-month, 2-year, and 10-year yields are uncorrelated with volatility.

Figure 3: Variation in Swaption Prices Explained by Yield Correlation
Swaption price can be uniquely composed as $P_t = P_t^y + P_t^{RV}$ where P_t^y is perfectly locallay correlated with the 6-month, 2-year, and 10-year yield and P_t^{RV} is uncorrelated with the same yields. The figure plots

$$F = \frac{\text{var}(P_t^{RV})}{\text{var}P_t} = 1 - \frac{\text{var}(P_t^y)}{\text{var}P_t}$$

for different price series. The blue line shows a very low correlation between the price changes of a swaption and the residual variance. The green line indicates a moderate correlation bewteen the time series of at-the-money swaption prices where the strike is updated by changes in the yield curve. The red line indicates a swaption straddle has nearly the same correlation as the at-the-money swaption prices with continually updated moneyness.
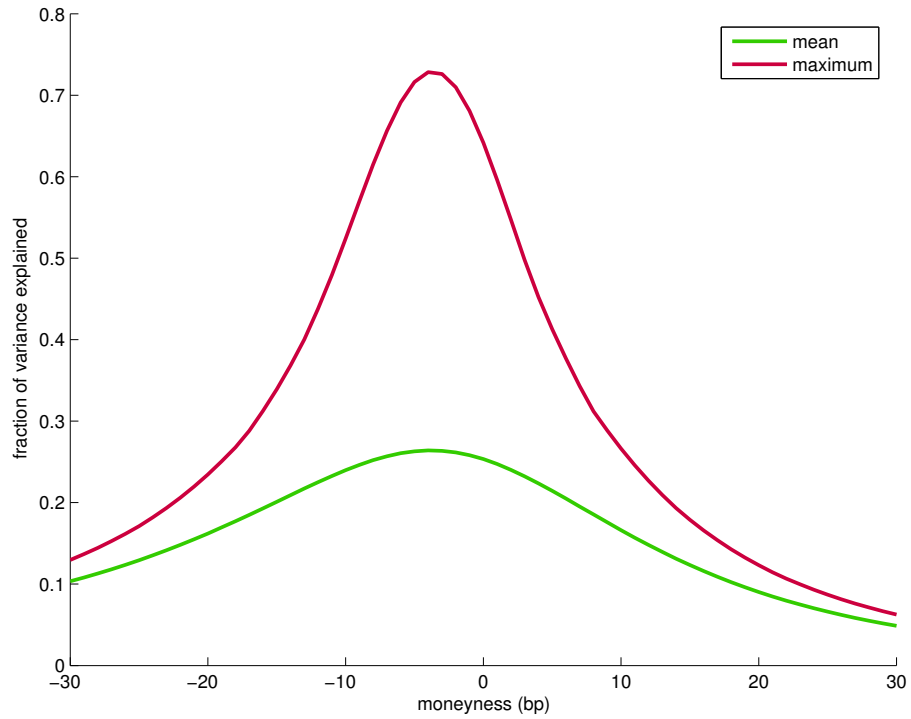
22

Figure 4: Reduction in Swaption Variance by Yield Correlation
This figure plots the mean of the model fraction of variance explained for a swaption straddle due to correlation with yields for the $A_1(4)$ model. As the straddle moves away from the money, the straddle become much less sensitive to residual variance risk.

5 Pricing Yield Risk

Fama and Bliss (1987), Campbell and Shiller (1991), and others have suggested that the shape of the yield curve drives risk premia that investors demand for holding long maturity bonds over short maturity bonds. A number of results (Cochrane and Piazzesi (2008), Duffee (2010) , Joslin et al. (2010a) , Ludvigson and Ng (2000), Rudebusch et al. (2006), Wright and Zhou (2009),and others) suggest that factors which have little incremental impact to the shape of the yield curve may be important for predicting excess returns for holding long maturity bonds. Indeed, Ludvigson and Ng (2000) find evidence that real and inflation risk factors have the ability to predict variation in bond excess returns above and beyond the level of interest rates. This raises the question of whether fixed income derivatives may be useful for identifying time-series variation in expected excess returns for holding long maturity bonds.

The theoretical possibility that volatility may incrementally forecast bond returns bond returns can be seen as follows. The (local) risk premium for exposure to a risk factor F_t is given by

$$\text{risk premium} = \mu_F^{\mathbb{P}}(X_t) - \mu_F^{\mathbb{Q}}(X_t).$$

That is, the expected excess returns for exposure to a risk is determined by the difference between the $\mathbb{P}$ and $\mathbb{Q}$ expected changes in the risk factor. We can reparameterize the model so that rather than the latent state variable X_t, we have the state variable $Y_t = (\ell_t, s_t, c_t, V_t^R)$ as in Section 4. In these terms, we can transform the model to be given in terms of Y_t which will have

$$\mu_Y^{\mathbb{Q}} = K_{0Y}^{\mathbb{Q}} + K_{1Y}^{\mathbb{Q}} Y_t, \tag{10}$$

$$\mu_Y^{\mathbb{P}} = K_{0Y}^{\mathbb{P}} + K_{1Y}^{\mathbb{P}} Y_t. \tag{11}$$

As seen in Figure 1, volatility has little incremental impact on bond yields relative to level, slope and curvature. Moreover, as elaborated in Section 7, volatility induces only a very small amount of variation in the convexity effect across maturities. Together these observations imply that volatility has little incremental impact on $\mathbb{Q}$-forecasts of of the level, slope and curvature. In this case (where volatility does not effect the $\mathbb{Q}$-forecasts of the yield factors), volatility will be useful for forecasting bond returns whenever volatility is incrementally informative for predicting future yields through $\mu_Y^{\mathbb{P}}$.

I decompose the risk premia into a component associated with the yield curve and a component due to the residual variance below. A one-year

standard deviation increase in the level of residual variance results in a decrease in expected return of approximately 1%. The variation in risk premia due to residual variance account for approximately 40% of the total variation. Although not the dominant term, the residual variance drives an economically meaningful portion of the risk premium.

This relationship can be borne in particular during the period from June 2004 to June 2006. During this period, the Federal Open Market Committee raised the target fed funds rate 25 basis points for 17 consecutive meetings. This period has been referred to as an a conundrum by then Fed Chairman Alan Greenspan because during this period the long rate remained relatively constant despite the increasing short. This conundrum is resolved either through changing expectations of future short rates (i.e. long rates could remain unchanged if investors anticipated the future Fed policy actions) or though declining term premiums. The model estimates indicates that this flattening of the yield curve was largely associated with declining risk premia. Thus, we can associate the flattening of the yield curve with a decline in risk premia for holding long maturity bonds (Section 7 further discusses the decomposition of the yield curve into expectations, term premia, and convexity effects). Figure 5 plots the slope of the yield curve (10 year-6 month rate) on the left axis and the implied volatility of an in 1 year-for 5 year swaption.[8] Here it is evident that the period is allows associated with a decline in yield volatility (perhaps due to increased transparency of monetary policy, as some suggest). These observation support the empirical results that volatility is an important determinant of the risk premium demanded for holding long maturity bonds.

6 Pricing Volatility Risk

Section 5 shows that volatility drives the risk premium that agents demand for holding long maturity bonds. We can consider also the converse question of what drives risk premia for holding volatility risk. That is, to what extent is the compensation agents demand for exposure to volatility risk is time-varying and can it be inferred from the prices of securities?

Again, considering the state vector $Y_t = (\ell_t, s_t, c_t, V_t^R)$ we can decompose the risk premium associated with exposure to the V_t^R risk factor. Figure 4

---

[8]These results partially reflect the fact that swaption implied volatility is quoted on a log of yields scale, rather than on a level of yields scale.
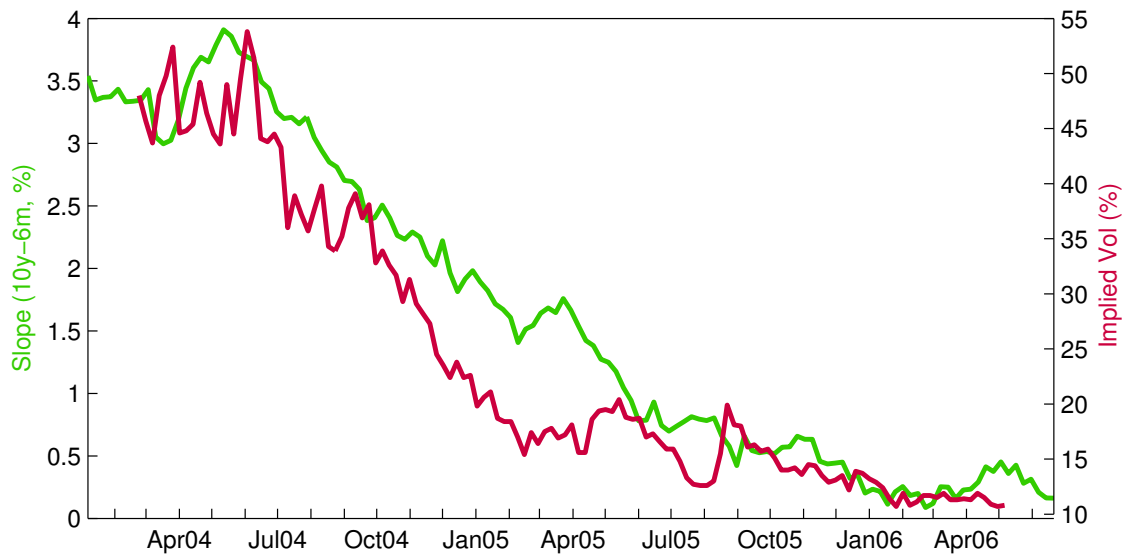
Figure 5: Long-Yield Conundrum

The figure plots the slope of the yield curve (10-year swap-implied zero rate minus the 6-month LIBOR rate) on the left axis and the implied volatility of the in 6-month-for-2-year swaption on the right axis. During this period, the yield curve became flat as the Fed continually raised interest rate. Also, implied volatilities declined similarly.

plots the time series of in 1 year for 5 year swaption implied volatilities. Also plotted are the model implied expected values of the payoffs of the options (that is, the $\mathbb{P}$-expected value rather than the $\mathbb{Q}$-expected value). Generally, the options have a negative expected return suggesting the option as an insurance premium, similar to the equities market (see, for example, Coval and Shumway (2001) or Pan (2002)). Also plotted are the expected value under the measure where the variance risk premium depends only on the slope of the yield curve.[9] This analysis shows that variation in the slope of the yield curve explains a portion of the variation in excess returns for holding variance risk

The results of Section 5 also suggest that it is difficult to be exposed directly to pure volatility risk over a long horizons and that volatility risk can be partially hedge by taking advantage of the moderate correlation of volatility risk with bond risk. Taken together, these results seem to indicate volatility risk may not be important. However, we can consider a synthetic security whose payoff is the realization of the future residual variance. Figure 6 plots the time series of the 1-year Sharpe ratio of such a security written on the 3-month conditional variance of a 5-year zero coupon bond with expiration in 1 year. Since the level of the residual variance is not identified, the exact payoff of the security at expiration is taken to be $\beta \cdot X_T$ where β is the loading defined in the residual variance at initiation. The Sharpe ratio varies through time with an average level of about 0.3. The Sharpe ratio tends to rise both in the Russian default and also in the 2001 recession. This suggest that the residual variance either is a risk that agents directly care about in their consumption decisions or at least is correlated with such a risk.

7 Role of Convexity in Bond Pricing

Long maturity bond yields reflect expectations of future interest rates, risk premia, and convexity effects. Fixing (risk-neutral) forecasts of future interest rates, convexity affects long maturity bond yields through Jensen's inequality as

$$\exp\left(E^{\mathbb{Q}}\left[-\int_0^T r_\tau d\tau\right]\right) < E^{\mathbb{Q}}\left[\exp\left(-\int_0^T r_\tau d\tau\right)\right].$$

---

[9]Formally, it the expectation computed assuming the dynamics follow the measure $\hat{\mathbb{P}}$ which differs from the $\mathbb{Q}$-measure by setting the drift of V_t^R equal to its $\mathbb{Q}$ drift except for the coefficient on the slope which maintains the coefficient under $\mathbb{P}$

Figure 6: Price of Variance Risk

This figure plots the sharpe ratio of the residual variance, defined as the risk which and is locally uncorrelated with the 6-month, 2-year, and 10-year yields.

The size of the convexity affect will be determined by the volatility of interest rates. I now turn to analyze the relative importance of this channel and its impact on bond prices and modeling the yield curve.

The T-year zero coupon yield can be decomposed in terms of an expectations effect $(y_{t,E})$, a risk premium $(y_{t,RP})$ and a convexity effect $(y_{t,C})$ as

$$y_t^T = y_{t,E}^T + y_{t,RP}^T + y_{t,C}^T ,\tag{12}$$

where

$$y_{t,E}^T \equiv \frac{1}{T} \int_t^{t+T} E_t^{\mathbb{P}} [r_\tau] \, d\tau ,$$

$$y_{t,RP}^T \equiv \frac{1}{T} \int_t^{t+T} (E_t^{\mathbb{Q}}[r_\tau] - E_t^{\mathbb{P}} [r_\tau]) d\tau ,\tag{13}$$

$$y_{t,C}^T \equiv -\frac{1}{T} \left(\log E_t^{\mathbb{Q}} \left[e^{-\int_t^{t+T} r_\tau d\tau} \right] + \int_t^{t+T} E_t^{\mathbb{Q}} [r_\tau] \, d\tau \right) .$$

The expectations term represents the bond price discounting with a yield to maturity equal to the average expected future short rate.

Figure 7 plots the decomposition of the 10-year zero coupon yield in terms of expectations, risk premia, and convexity effects as defined above. Each of the terms are an affine function of the state variable whose loading can be computed by solving a Riccati differential equation or linear constant coefficient ordinary differential equation. The figure shows that the variation in yields are dominated by expectations and risk premium effects and that the convexity effects are quite small (see also Gupta and Subrahmanyam (2000)).

Table 5 gives the magnitude of the convexity effect across maturities for the various models specifications. For comparison, an $A_1(3)$ model, estimated on the same data but not inverting the swaption, is added to both tables. Panel A shows that the average convexity effects are small for the 2-year zero coupon bond, around one basis points. Extending the maturity to ten years, the average size of the convexity effects becomes more economically meaningful, reaching around 15 basis points. However, Panel B shows that although the 10-year convexity effect is larger, the variation is still quite small with a weekly standard deviation of less than a basis point. Extending to thirty year bonds, the average convexity effect becomes quite important, but the variation still remains quite small. The reason behind this is quite intuitive – over short horizons convexity effects are generally unimportant, while over long horizons mean reversion in the level of volatility implies that
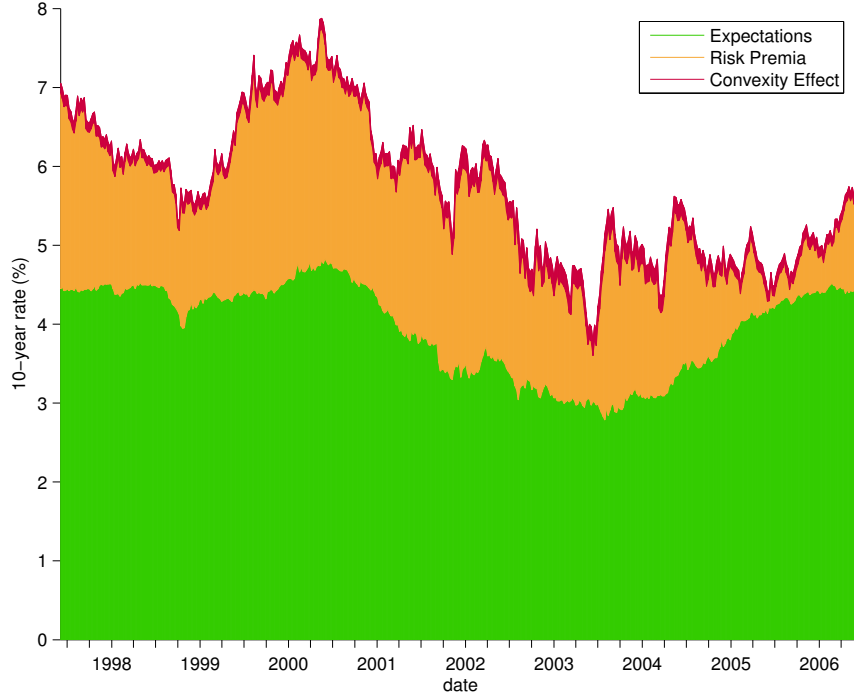
Figure 7: Yield Decomposition
This figure plots model-implied decomposition of the 10-year yield into expecations, risk premium, and convexity components for the estimated $A_1(4)$ model.

$$y_{\text{expectation}} = \frac{1}{10} \int_t^{t+10} E_t^P[r_\tau]d\tau \,,$$

$$y_{\text{risk premia}} = \frac{1}{10} \int_t^{t+10} (E_t^Q[r_\tau] - E_t^P[r_\tau])d\tau \,,$$

$$y_{t,\text{convexity}} = -\frac{1}{10}\Big(\log E_t[e^{- \int_t^{t+10} r_\tau d\tau}] + \int_t^{t+10} E_t^Q[r_\tau]d\tau \Big) \,.$$

the current level of volatility has only a small impact on the amount of convexity in long maturity yields.

	$A_1(3)$	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
2 year	0.48	0.46	0.79	0.46	1.80
5 year	4.06	4.01	4.78	4.02	5.20
10 Year	16.19	15.65	17.79	15.83	15.00
30 Year	85.48	79.15	168.29	80.06	90.95

(a) Average Convexity Effects

	$A_1(3)$	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
2 year	0.02	0.03	0.07	0.04	0.11
5 year	0.13	0.28	0.36	0.27	0.28
10 Year	0.39	0.84	1.01	0.83	0.38
30 Year	0.80	1.84	2.69	1.82	0.27

(b) Time Variation of Convexity Effects

Table 5: Convexity Effects in Zero Coupon Bond Yields
The top panel gives the average model-implied convexity effect in basis points for different bond maturities. The bottom panel gives the sample standard deviation, in basis points, of weekly changes in the model-implied convexity effects. The convexity effect of an T-year yield:

$$C_t(T) \equiv \frac{1}{T} \log E[e^{-\int_t^{t+T} r_\tau d\tau}] + \frac{1}{T} E^Q[\int_t^{t+T} r_\tau d\tau]$$

All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed.

These results show that convexity plays only a small role in bond prices. The fact that convexity effects are small implies that a dynamic term structure model may exhibit arbitrary correlation between the first few principal components and volatility under very parsimonious conditions. For example, consider the $A_1(4)$ model and approximate (6) by eliminating the quadratic convexity term,

$$\dot{B} \approx -\rho_1 + (K_1^{\mathbb{Q}})'B \,. \tag{14}$$

If there is a risk factor which affects the conditional volatility of yields, but does not affect risk-neutral expectations of future rates, volatility will only be related to the yield curve through the small convexity effect and correlation between the risk factors.[10]

As Collin-Dufresne et al. (2009) argue, the fact that convexity effects are small suggests that volatility may be poorly identified from the cross section of bond prices.[11] Indeed, even in a model where the constraints for USV are strongly violated, volatility may only be identified through the convexity effect in a very sensitive manner. More precisely, although volatility may be directly inferred from bond prices, it is only through solving the numerically unstable equation $Ax = b$ where A is nearly singular. The near singularity of A means that small errors, for example measurement errors in the yields or estimation errors, may result in large errors in the inferred volatility. For example, if 6-month, 2-year, 5-year, and 10-year zero coupon yield are used to infer volatility, the unconstrained $A_1(4)$ estimates indicate that the matrix A will have a very high condition number (6,299 – condition numbers over 30 suggest multicolinearity) and thus nearly non-singular.

To highlight the mechanism, define a model with a factor that has volatility *unspanned by expectations* by

Defintion 7.1. *a diffusive $\mathbb{Q}$-short rate model,*

$$dr_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t$$

has volatility unspanned by expectations if there exists a change of variable $Y_t = f(X_t)$ such that Y_t^1 drives volatility but does not affect $\mathbb{Q}$-expectations of future interest rates ($\partial\mu/\partial y^1 \equiv 0$).

---

[10]More formally, the precise condition is the existence of an eigenvector of $K_1^{\mathbb{Q}}$ which is orthogonal to ρ_1 and loads on the volatility factors.

[11] Andersen and Benzoni (2010) stress the theoretical deficiency of general affine models to produce low correlation between volatility changes and yield changes.

A non-degenerate affine term structure model has volatility unspanned by expectations if there exists an eigenvector of K_1 which is orthogonal to ρ_1. However, not all classes of term structure models admit volatility risk factors unspanned by expectations. Gaussian affine term structure models clearly cannot admit weakly spanned volatility risk factors (the eigenvector condition implies the model will be degenerate). Additionally, it is easy to show that quadratic affine term structure models, as in Longstaff (1989), Ahn et al. (2002) , Li and Zhao (2006) and others, do not admit weakly spanned volatility risk factors.

The mechanism at work in models with unspanned volatility and expectations unspanned volatility are very different. In USV models, there exists small convexity effects whose differential effects across maturities must be exactly cancelled by a corresponding expectations effect. This requires restrictions on the number of stochastic convexity effects generated and the rates of mean reversions of the risk factors. For example, in the $A_1(3)$ specification, there are 10 parameter constraints required.[12] In contrast, when volatility is unspanned by expectations, it will by spanned by yields. However, this is only through the convexity effects which we have seen empirically show very small time series variation. Furthermore, this requires only a single one parameter constraint.

The simple condition, which ignores the small convexity effect, is very different from the conditions required for unspanned stochastic volatility, which explicitly cancels the convexity effect. For example, Table 6 shows both USV models have *two* persistent risk factors with long half-lives. This is because the convexity effect generated by a persistent risk factor with stochastic volatility can only be canceled by a risk factor with twice the rate of mean reversion. Table 7 shows that both a Lagrange-multiplier test using the restricted estimates and a Wald test using the unrestricted estimates reject the restriction on the rates of mean reversion. Additionally, a likelihood ratio test of the constrained USV model against the unconstrained model strongly rejects the USV restrictions for both the $A_1(4)$ and $A_2(4)$ models. The economic effect of the mean reversion restrictions can also be understood by comparing the ability of the models to price 30-year zero coupon bonds

---

[12] Joslin (2006) shows that in order for the convexity effect to cancel, three types of restrictions must hold: (1) some factor mean reversions must related in a 2:1 ratio in order to possibly cancel a quadratic convexity effect, (2) some factors must have constant volatility in order to not generate convexity effects, and (3) volatility must affect expectations of future rates in exactly the right way to cancel the convexity effect.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
1st Eigenvalue	1.492	1.233	1.563*	3.833*
2nd Eigenvalue	0.642	0.244*	0.624	1.628*
3rd Eigenvalue	0.317*	0.067	0.348*	0.237
4th Eigenvalue	0.050	0.034	0.053	0.118

Table 6: Eigenvalues Under The Risk Neutral Measure

Eigenvalues of the mean reversion matrix, κ^Q, under the risk-neutral measure. Asterisks denote the eigenvalues of the CIR factors. All models were estimated on weekly data for the period from June 4, 1997 to June 21, 2006. The USV superscript denotes an affine model with USV constraints imposed.

in Table 3. These bonds were not used in estimation and thus represent an out-of-sample comparison of the models. For the longer maturities, the non-USV models price the yields reasonably well with root mean square errors ranging from 10 to 17 basis points. The errors for the USV models are much larger. The cause for the larger error can be attributed to the restrictions on the rates of factor mean reversion imposed by USV. Consistent with the rejection of the restriction of two persistent risk factors, show that the USV models have very large errors in pricing the long maturities bonds.

Table 6 shows the eigenvalues under the risk neutral measure of the drift feedback matrix, $\kappa^{\mathbb{Q}} \equiv -K_1^{\mathbb{Q}}$. The eigenvalues determine the level of persistent of shocks to the risk factors. An eigenvalue of λ corresponding to a half-life of $\log(2)/\lambda$. For each model, there is at least one very persistent "level" factor. In the case of the USV models, there is also a second factor with twice the rate of mean reversion to cancel the convexity effect generated by the stochastic volatility of the most persistent factor. For example, in the $A_1(4)^{USV}$ model, there are eigenvalues of .034 and .067, corresponding to half-lives of 20.4 and 10.2 years, respectively. This condition of two persistent factors results in a misspecification at the long end of the curve and larger mispricings.

	$A_1(4)$	$A_2(4)$
Lagrange Multiplier (mean reversion)	6.31 (.006)	14.73 ($\ll$ .001)
Likelihood Ratio (full model)	2,694 ($\ll$ .001)	4804 ($\ll$ .001)

Table 7: Statistical Tests of USV constraints

This table gives the test statistics for the constraints required for the $A_1(4)$ and $A_2(4)$ models to exhibit unspanned stochastic volatility. The mean-reversion constraint is tested by a Lagrange multiplier test. In addition, the complete set of restrictions are rejected by a likelihood ratio test.

8 Conclusion

In this paper, I show that when the covariance structure of risk factors and market prices of risk are not restricted, low-dimensional dynamic term structure models are able to simultaneously capture the price dynamics in bond and bond option markets. I show that under parsimonious conditions there can exist a residual component of volatility risk largely uncorrelated with yield changes. This residaul volatility risk is an important determinant of the risk premium that investors demand for holding long maturity bonds. Conversely, the shape of the yield curve is related to the premium that investors demand for bearing interest rate volatility risk. In my estimation, I find empirical evidence rejecting conditions for unspanned volatility. Instead a component of volatility has very little effect on the cross-section of bond yields due to having a small effect on risk neutral expectations of future short rates and inducing little variation in convexity effects. Finally, I develop computational methods for pricing options and extend the technique to provide maximum likelihood estimation of general affine diffusions which can be used in a number of contexts.

A   Model Specification

For the affine term structure model

$$r_t = \rho_0 + \rho_1 \cdot X_t,$$
$$dX_t = \mu_t dt + \sigma_t dB_t,$$
$$\mu_t = K_0 + K_1 X_t,$$
$$\sigma_t \sigma_t' = H_0 + H_1 \cdot X_t,$$

where

$$K_1 = \begin{bmatrix} K_V & 0 \\ K_{VG} & K_G \end{bmatrix}, \quad H_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_0^G \end{bmatrix}, \quad H_1^i = \begin{bmatrix} \Sigma_i^V & 0 \\ 0 & \Sigma_i^G \end{bmatrix},$$

with K_V an $M \times M$ matrix, K_G and Σ_i^G $(N-M) \times (N-M)$ matrices, and Σ_i^V $M \times M$ matrices, The *drift normalized* canonical representation as follows. For the parameters $\Theta = (\rho_0, \rho_1, K_0^P, K_1^P, K_0^Q, K_1^Q, C_{M+1}^G, C_{M+2}^G, \ldots, C_N^G)$, impose the constraints:

1. K_G^Q is diagonal with entries increasing on the diagonal.

2. C_i^G is lower triangular and gives the Cholesky factorization of Σ_i^G: $\Sigma_i^G = (C_i^G)^T (C_i^G)$.

3. $K_{0,n}^Q = 0$, $n > M$.

4. $\Sigma_{i,jk}^V = 1$, if $j = k = i$, or 0 otherwise.

5. $\rho_{1,n} = 1$, $n > M$.

6. $\rho_{1,n} < \rho_{1,n+1}$, $n < M$.

7. $K_{V,ij}^P > 0$, $K_{V,ij}^Q > 0$ if $i \neq j$.

8. $K_{0,n}^P \geq \frac{1}{2}$, $K_{0,n}^Q \geq \frac{1}{2}$, $n \leq M$.

B Pricing

This appendix presents a computationally efficient method for computing the transform given in Duffie et al. (2000):

$$G(y) = E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T}\{\delta \cdot X_T \le y\}],$$
$$\hat{G}(y) = E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T + i\delta \cdot X_T}], \tag{15}$$
$$G(y) = \frac{\hat{G}(0)}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{t}\mathrm{Im}(\hat{G}(t)e^{ity})dt.$$

In computing the transform, we can use the fact that $Y = \delta \cdot X_P$ is roughly normally distributed under the forward measure,F, where

$$\frac{dF}{dQ} = \frac{e^{-\int_0^P r_\tau d\tau + d \cdot X_T}}{E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T}]}. \tag{16}$$

More precisely, $\hat{G}(t) \approx ce^{-\sigma_Y^2 t^2/2 + it\mu_Y}.$[13] In the case of an $A_0(N)$ Gaussian model, this equation is exact. Considering this case for now, the Levy integral then becomes:

$$
\begin{aligned}
I &= \int_0^\infty \frac{1}{t}\mathrm{Im}(\hat{f}(t)e^{-ity})dt \\
&= \int_0^\infty \frac{1}{t}\mathrm{Im}(e^{-\frac{\sigma^2 t^2}{2}}e^{it\mu}e^{-ity})dt \\
&= \int_0^\infty \frac{\sin(t(\mu - y))}{t}e^{-\frac{\sigma^2 t^2}{2}}dt \\
&= \int_0^\infty g(t)w(t)dt
\end{aligned}
$$

Where $w(t) = e^{-\sigma^2 t^2/2}, g(t) = \sin((\mu - y)t)/t$. $w(x)$ is a scaling of the weighting function e^{-t^2} used in Gauss-Hermite quadrature. By using flexibility in both the choice of nodes and weights, Gauss-Hermite quadrature allows very accurate computations for integrals of the form $\int g(t)e^{-t^2}dt$ with very few nodes. This suggests that, after appropriate scaling, Gauss-Hermite quadrature will be an accurate way to compute the inversion integral.

---

[13]It is important to note that a smooth density function implies fast decay of the Fourier transform.

In general, we can write the Levy integral as:

$$I = \int_0^\infty \underbrace{\frac{1}{t}\text{Im}(\hat{f}(t)e^{-ity}e^{\sigma^2 t^2})}_{g(t)} e^{-\sigma^2 t^2} dt$$

$$\approx \sum_i g(t_{i,\sigma})w_{i,\sigma}$$

Two points also become clear:

1. **Scale Matters.** If we are computing the transform integral, we must integrate on approximately $t \in [-\frac{3}{\sigma}, \frac{3}{\sigma}]$ before rescaling. This means if we are directly computing this integral and we are using options with various maturities (so that σ will vary) *any* quadrature scheme must take this into account.

2. **Out of moneyness increases oscillation of integrand** By rescaling to change the integral to:

$$I = \int_0^\infty \sigma \frac{\sin(\frac{\mu-y}{\sigma}u)}{u} e^{-u^2/2} du$$

we see that the integral will have a weighting function times a decaying oscillatory terms. The frequency of oscillation increases as the we move more standard deviation for μ.

Example

We now turn to an example of computing forward probabilities. Consider the risk-free $A_1(2)$ term structure model in Duffie et al. (2003a). To emphasize the generality of the approach, I augment the model with jumps occurring with intensity $\lambda = 1$ of size $\pm 1.5\%$ in the short rate.

I then compute a term involved in pricing a zero coupon bond option:

$$E_0[e^{-\int_0^T r_s ds} e^{B(\tau)\cdot X_T}\{-\frac{A(\tau) + B(\tau)\cdot X_T}{\tau} \geq f_0 + m\}]$$

Here τ is the maturity of the underlying zero coupon bond (which has log price $A(\tau) + B(\tau) \cdot X_0$ when the state is X_0), T is the expiry of the option, and f_0 is the corresponding forward rate with m a moneyness adjustment. I

compute this term for an option on a 5-year zero coupon bond with expiry of 6 months ($T = .5, \tau = 5$.) The strikes are adjusted from the corresponding forward rate of 7.27%. The initial state was taken to be the long run mean, $X_0 = \theta_P$.

Figure 8 shows the integrand (scaled by σ) in the Levy inversion integral for the various strikes. In each case the integrand can be seen to be $e^{-x^2}g(t)$ where $g(t)$ is a decaying oscillatory function which is more oscillatory the more the option is out of the money. The left panels plot the integrand itself, where the right panel plots $g(t)$. The figure also plots in red the nodes used for the quadrature with $n = 5$ nodes.

Another method of computing the forward probability $P(b \cdot X_T \leq y)$ would be to use a cumulant expansion for the random variable $b \cdot X_T$.[14] This amounts to doing a Taylor series expansion of the right hand panel. As can be seen, when the option is near the money, the Taylor series will be accurate, except for large values of t which are given little weight in the integral. However, as the options becomes more out of the money (the lower panel), the Taylor series approximation will become inaccurate. In contrast, the quadrature scheme is able to both pick up the oscillatory nature of the integrand and focus on the region which is important for the integral.

Table 8 reports the accuracy of the quadrature for various number of nodes. The reference value was computed using Simpson's rule with 10000 nodes spaced on the interval $[0, 6/\sigma]$. The variable n_σ measures how many standard deviation (under the forward measure) the option is out of the money. For a fixed number of nodes, the accuracy decays as the option goes out of the money since the Levy integrand becomes more oscillatory. For options which are within a standard deviation of the being at the money, the quadrature scheme is quite accurate with even just 3 nodes.

C Computing Exact Likelihood

Because the conditional characteristic function is known in terms of the solution to an ordinary differential equation, the transition likelihood for an

---

[14]The cumulants will be affine in the state and can be obtained by repeatedly differentiating the original Riccati equation. In the case of a forward measure, the cumulants can again be computed by differentiating a different Riccati equation.

affine diffusion can be recovered from the characteristic function:

$$\hat{f}(s|x_t) = E[e^{is \cdot X_{t+1}}|X_t = x_t]$$

$$f(x_{t+1}|x_t) = \frac{1}{(2\pi)^n} \int \hat{f}(s|x_t) e^{-ix_{t+1} \cdot s} ds \quad (17)$$

However, direct computation of this integral is often intractable. Two ideas are used in order to simplify the computations involved. First, the integrand in the inverse Fourier transform of a transition becomes more oscillatory as the transition varies from the expected transition. In order to remove the oscillations, a transition measure is defined where the observed transition becomes the a likely transition. That is,

$$\hat{f}(s) \approx e^{i\mu_t \cdot s - \frac{1}{2}s^\top \Sigma_t s}$$

$$\text{Real}(\hat{f}(s)e^{-is \cdot x_{t+1}}) \approx \text{Real}(e^{i(\mu_t - x_{t+1}) \cdot s - \frac{1}{2}s^\top \Sigma_t s})$$

$$= \cos(s \cdot (\mu_t - x_{t+1}))e^{-\frac{1}{2}s^\top \Sigma_t s}$$

If we define $dT/dP = e^{a \cdot X_{t+1}}/E_t[e^{a \cdot X_{t+1}}]$, under T, $E_t[X_{t+1}] \approx \mu + \Sigma a$.[15] So, by choosing $a = \Sigma^{-1}(x_{t+1} - \mu)$, $E_t^T[X_{t+1}] \approx x_{t+1}$ and

$$f(x_{t+1}|x_t) = f^T(x_{t+1}|x_t) \times \frac{dP}{dT}(x_{t+1})$$

After this change of measure, $\hat{f}^T \approx e^{ix_{t+1} \cdot s - \frac{1}{2}s^\top \Sigma_t s}$ and so the integrand in the inverse Fourier transform to compute $f^T(x_{t+1})$ is approximately $e^{-\frac{1}{2}s^\top \Sigma s}$. Thus the integral

$$f^T(x_{t+1}) = \frac{1}{(2\pi)^n} \int \hat{f}^T(s)e^{-is \cdot x_{t+1}} ds$$

$$= \frac{1}{(2\pi)^n} \int w(s)e^{-\frac{1}{2}s^\top \Sigma s} ds$$

---

[15] Note that when there are CIR factors we must consider that $E_t[e^{a \cdot X_\tau}]$ is finite for all τ only when a is in the domain of attraction of the fixed point of the affine differential equation. However, even when this is not the case the expectation will be finite for τ small and calculation show this range is reasonably large when boundary non-attainment is enforced.

Strike	Price	n_σ	3 nodes	5 nodes	8 nodes
7.27%	30.55	0.03	0.00029	1.02e-006	8.47e-009
8.27%	6.60	1.24	0.00409	2.18e-005	3e-008
9.27%	0.55	2.46	0.178	0.00247	7.49e-006

Table 8: Accuracy of Quadrature
This table presents the accuracy of the adaptive Gauss-Hermite quadrature scheme for computing option prices. The first column gives the strike of the 6-month option on 5-year zero coupon bond. The exact price of the option is given in the second column and the number of standard deviation (under the risk neutral measure) the option is out of the money is given in the third column. The remaining columns give the accuracy of the quadrature scheme.

where $w(s) = e^{-\frac{1}{s}^\top \Sigma s}/\hat{f}^T(s)e^{-is\cdot x_{t+1}}ds \approx 1$. This multidimensional integral is suitable to be evaluated by Gauss-Hermite quadrature with very few nodes. Also, since the integrand is odd, only half of the evaluations need actually be done. Though the curse of dimensionality is still present, the computation now become tractable since even 4 nodes gives reasonable accuracy. Evaluating the inverse transform for rare transition with highly oscillatory inverse Fourier transform integrands would require solving hundreds of millions of differential equations (4^4 versus 100^4, for example).
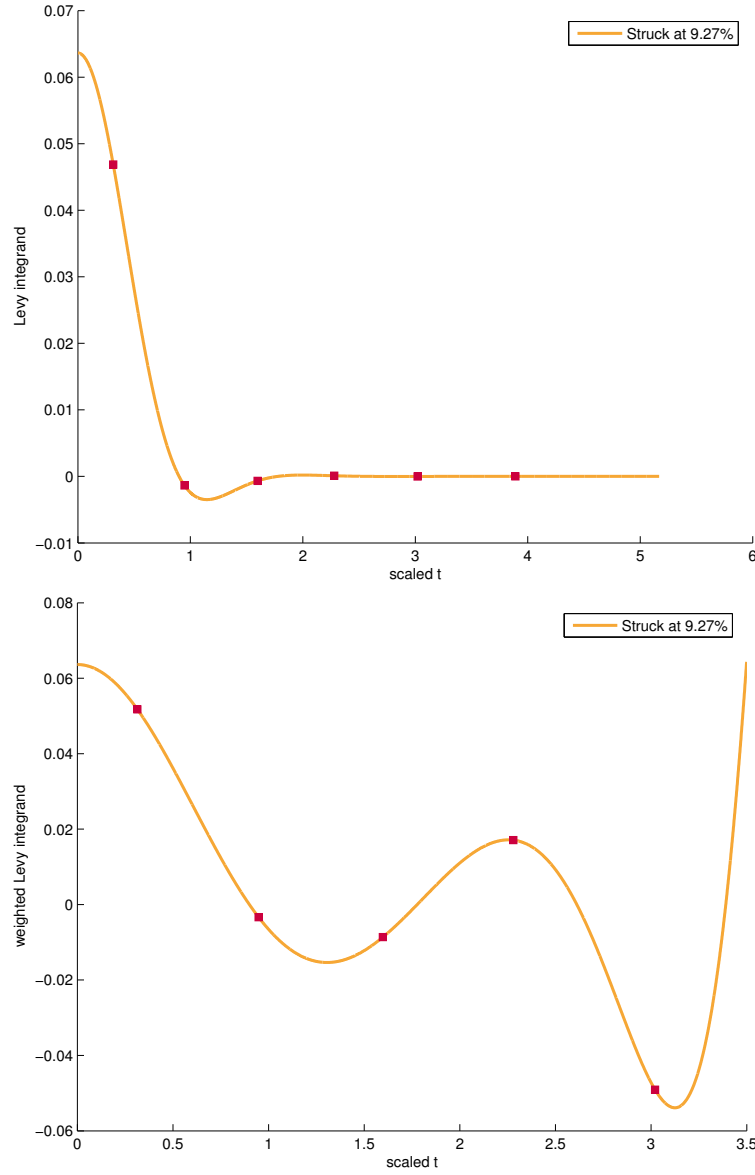
Figure 8: Levy Integrand

The top panel plots the Levy-integrand used in computing the value an out-of-the money bond option. The integrand is approximately a scaled normal density times an oscillatory function. The bottom panel plots the oscillatory multiplier by weighting the integrad. The squares indicate nodes used in Gauss-Hermite quadrature.

References

Dong-Hyun Ahn, Robert Dittmar, and Ronald Gallant. Quadratic term structure models: Theory and evidence. *Journal of Financial Economics*, 15(1):243–288, 2002.

Yacine Ait-Sahalia. Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance*, LIV(3):1361–1395, 1999.

Yacine Ait-Sahalia and Robert Kimmel. Estimating affine multifactor term structure models using closed-form likelihood expansions. Journal of Financial Economics, 98:113–144, 2010.

Torben Andersen and Luca Benzoni. Do bonds span volatility risk in the u.s. treasury market? a specification test for affine term structure models. *forthcoming, Journal of Finance*, 2010.

Michael W. Brandt and Pedro Santa-Clara. Simulated likelihood estimation of diusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, 63:161–210, 2002.

Wolfgang Buhler, Marliese Uhrig-Homburg, Ulrich Walter, and Thomas Weber. An empirical comparison of forward-rate and spot rate models for valuing interest-rate options. *Journal of Finance*, LIV(1):269–305, February 1999.

John Y. Campbell and Robert J. Shiller. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58(3):495–514, May 1991.

Marine Carrasco, Mikhail Chernov, Jean-Pierre Florens, and Eric Ghysels. Efficient estimation of jump diffusions and general dynamic models with a continuum of moment conditions. *Journal of Econometrics*, 2006.

Patrick Cheridito, Damir Filipovic, and Robert L. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83(1):123–170, January 2007.

John Cochrane and Monika Piazzesi. Decomposing the yield curve. Working Paper, 2008.

P. Collin-Dufresne, R. Goldstein, and C. Jones. Identification of maximal affine term structure models. Journal of Finance, 63(2):743–795, 2008.

Pierre Collin-Dufresne and Robert S. Goldstein. Pricing swaptions within an affine framework. *Journal of Derivatives*, 10(1):1–18, Fall 2002a.

Pierre Collin-Dufresne and Robert S. Goldstein. Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *Journal of Finance*, 57(4):1685–1730, August 2002b.

Pierre Collin-Dufresne, Robert S. Goldstein, and Christopher S. Jones. Can the volatility of interest rates be extracted from the cross section of bond yields? an investigation of unspanned stochastic volatility. *Journal of Financial Economics*, 94(1):47–66, 2009.

Josh Coval and Tyler Shumway. Expected option returns. Journal of Finance, pages 983–1009, 2001.

Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978, October 2000.

Qiang Dai and Kenneth J. Singleton. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, 63(3):415–441, March 2002.

Qiang Dai and Kenneth J. Singleton. Term structure dynamics in theory and reality. *Review of Financial Studies*, 16(3):631–678, Fall 2003.

J. Dreissen, P. Klaasen, and B. Melenberg. The performance of multi-factor term structure models for the pricing and hedging caps and swaptions. *Journal of Financial and Quantitative Analysis*, 38(2):635–672, 2003.

G. Duffee. Information in (and not in) the term structure. 2010.

Gregory R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):405–443, February 2002.

D. Duffie. *Dynamic Asset Pricing Theory, 3rd edition*. Princeton University Press, Princeton, N.J., 2001.

Darrell Duffie, Lasse Pedersen, and Kenneth J. Singleton. Modeling sovereign yield spreads: A case study of russian debt. *Journal of Finance*, 58(1): 119–160, 2003a.

J. Darrell Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, October 1996.

J. Darrell Duffie and Kenneth J. Singleton. Simulated moments estimation of markov models of asset price. *Econometrica*, 61:929–952, 1993.

J. Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, November 2000.

J. Darrell Duffie, Damir Filipovic, and Walter Schachermayer. Affine processes and application in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003b.

Eugene F. Fama and Robert R. Bliss. The information in long-maturity forward rates. *American Economic Review*, 77(4):680–692, September 1987.

Anurag Gupta and Marti Subrahmanyam. An empirical examination of the convexity bias in the pricing of interest rate swaps. Journal of Financial Economics, 55:239–279, February 2000.

Ravi Jagannathan, Andrew Kaplin, and Steve Sun. An evaluation of multi-factor cir models using libor, swap rates, and cap and swaption prices. *Journal of Econometrics*, 116(1-2):113–146, September-October 2003.

S. Joslin, M. Priebsch, and K. Singleton. Risk premiums in dynamic term structure models with unspanned macro risks. 2010a.

S. Joslin, K. Singleton, and H. Zhu. A new perspective on gaussian dynamic term structure models. *Review of Financial Studies*, 2010b. forthcoming.

Scott Joslin. Can unspanned stochastic volatility models explain the cross section of bond volatilities? *Working Paper, Stanford University*, 2006. Working Paper, Stanford University.

Haitao Li and Feng Zhao. Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. *Journal of Finance*, 61(1):341–378, 2006.

Robert Litterman and Jose Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1(1):54–61, June 1991.

Francis Longstaff. A nonlinear general equilibrium model of the term structure of interest rates. *Journal of Financial Economics*, 23:195–224, 1989.

Sydney C. Ludvigson and Serena Ng. Macro factors in bond risk premia. *The Review of Financial Studies*, 22(12):5027–5067, August 2000.

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

Glenn Rudebusch, Eric Swanson, and Tao Wu. The bond yield "conundrum" from a macro-finance perspective. *Monetary and Economic Studies*, 24 (S-1):83–128, December 2006.

Kenneth J. Singleton. Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102:111–141, 2001.

Kenneth J. Singleton and Len Umantsev. Pricing coupon-bond options and swaptions in affine term structure models. *Mathematical Finance*, 12: 427–446, 2003.

Anders Trolle and Eduardo S. Schwartz. A general stochastic volatility model for the pricing of interest rate derivatives. Review of Financial Studies, 22 (5):2007–2057, 2009.

J. Wright. Empirical evidence on term premiums from an international panel dataset. *American Economic Review*, 2009. forthcoming.

Jonathan Wright and Hao Zhou. Bond risk premia and realized jump volatility. *Journal of Banking and Finance*, 33:2333–2345, 2009.