

Pricing and Hedging Volatility Risk in Fixed Income Markets*

Scott Joslin[†]

Current Draft: September 30, 2007

*I am very grateful to my advisor Ken Singleton for numerous discussions and comments. I also appreciate comments provided by my dissertation committee, Darrell Duffie and Ilya Strebulaev, as well as Caio Almeida, Snehal Banerjee, Jeremy Graveline, Yaniv Konchitchki, Jun Pan, and seminar participants at Caltech, Chicago GSB, Georgia Institute of Technology, London School of Economics, McGill University, MIT Sloan, NYU Stern, Pennsylvania State University, Stanford GSB, UC Berkeley, Vanderbilt University, and Washington University.

[†]MIT Sloan School of Management. 50 Memorial Drive, Cambridge, MA 02142.
sjoslin@mit.edu

Abstract

Although financial theory predicts a strong link between bond yields and bond option prices, recent research has had limited success at pricing both bonds and bond options using low-dimensional dynamic term structure models. 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. Moreover, this model resolves the puzzle in the literature that a substantial component of the bond option volatility is largely uncorrelated with bond yields. The structure of risk in my model that gives rise to this feature of volatility is distinct from that inherent in recent models with “unspanned stochastic volatility.” In fact, the restrictions on the conditional distributions of the risk factors that lead to the derivative-specific factors in the latter models are strongly rejected by the data. Consistent with my model’s ability to price both bonds and options, I find that dynamic hedging strategies using bonds alone produce reasonably good hedges for derivative positions. Finally, I find that agents are risk averse to changes in volatility and that a component of this volatility risk is an important determinant of the risk premiums that agents demand for holding long maturity bonds.

1 Introduction

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 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). Moreover, model-free principal components analyses, e.g., Heidari and Wu (2003), show that the level, slope, and curvature term-structure factors explain only about 60% of the cross-sectional variation in option-implied volatilities.² This paper develops a dynamic term structure model with four risk factors that resolves both of these empirical puzzles: the model’s mean-squared pricing errors on options are less than the bid-ask spreads, and one of the risk factors drives option volatilities while being only weakly correlated with bond yields, leading to the ability of the model to price both bonds and bond options.

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 the affine process specification given in Joslin (2006) which allows for a richer covariance structure among risk factors than the specification of Dai and Singleton (2000). In contrast, Jagannathan et al. (2003) examine multi-factor Cox-Ingersoll-Ross models which fails to capture both the first- and second-moment properties of bond yields (e.g., Dai and Singleton (2000,2002)), and so would not be expected to accurately price options on bonds. Similarly, Buhler et al. (1999) consider only special cases of the general specification of Dai and Singleton (2000), requiring independent risk factors with restricted conditional first-moments. The covariance structure of the risk factors is critical both in the pricing of derivatives and in capturing risks that drive option implied volatilities but do not affect the level, slope, and curvature.

Importantly, the success of my model is achieved through a very different mechanism from that underlying models with “unspanned stochastic

¹Dai and Singleton (2003) for a survey of the dynamic term structure literature.

²Litterman and Scheinkman (1991) show that the first three principal components of bond yields have loadings that look roughly like level, slope, and curvature factors, and that these factors explain over 95% of the variation in these bond yields.

volatility (USV).” The finding that a large component of the variation in option implied volatilities is uncorrelated with the principal components of bond yields motivated Collin-Dufresne and Goldstein (2002b) to develop an arbitrage-free model with derivative-specific risk factors. Subsequently, Bikbov and Chernov (2005), Collin-Dufresne et al. (2006a), Han (2006), Li and Zhao (2006), and others have adopted models embodying USV to study the pricing of bond options. In order that there are unspanned risk factors which affect option prices but not bond prices, the convexity effect implicit in long maturity bond yields must be cancelled by an expectations effect. USV models impose precise conditions so that this convexity effect is exactly cancelled. However, as elaborated in Section 4, I show that a component of volatility risk can be uncorrelated with the level, slope, and curvature factors under more parsimonious, and different, conditions because the convexity effect is small. Using statistical tests and out-of-sample pricing, I find that the conditions required for cancellation of the convexity effect are rejected by the data. However, the general model, which strongly violates the USV restrictions, produces a component of volatility risk which is uncorrelated with the level, slope, and curvature factors.

The second feature of my analysis is the dependence of the market price of risk on the state of the economy. In particular, I follow Cheridito et al. (2006) in parameterizing the market prices of risk, and this allows for a flexible market price of volatility risk which is crucial for matching the option price dynamics. Both Buhler et al. (1999), and Jagannathan et al. (2003) examine a setting where the market price of risk is linear in the state of the economy. Duffee (2002) and Dai and Singleton (2002) show that such a simple market price of risk is inconsistent with the empirical violations of the expectations hypothesis. However, the market price of risk specification that they use is still very restrictive in how volatility risk is priced.

The goodness-of-fit of my model is assessed in various ways. Beyond standard likelihood ratio tests of constraints related to the structure of volatility risk, I explore the performance of model-based hedges of options. I find that the amount of correlation between level, slope, and curvature changes and yield volatility changes is stochastic and typically ranges between 70% to 80%. Consistent with this, and the results of Fan et al. (2001), I find that bond positions are able to hedge most of the risk in a swaption straddle position, which is particularly sensitive to volatility risk. I find that over long horizons a dynamic hedging strategy is required.

I also find that the component of volatility which is uncorrelated with

the first three principal components is an important factor in determining the risk premia associated with holding long maturity bonds, explaining approximately 50% of the variation in expected excess returns. Additionally, I find that investors are risk averse to movements in this residual component of volatility.

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. Section 4 examines the role of convexity in bond prices. Hedging analysis is carried out in Section 5. The pricing of volatility risk is examined in Section 7. Finally, Section 9 concludes.

2 Model and Estimation Strategy

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

$$\begin{aligned} r_t &= \rho_0 + \rho_1 \cdot X_t, \\ dX_t &= \mu_t dt + \sigma_t dB_t^P, \\ \mu_t &= K_0 + K_1 X_t, \end{aligned} \tag{1}$$

I consider $A_1(4)$ and $A_2(4)$ models where either one or two factors drive volatility. For example, in the $A_2(4)$ case, $\sigma_t \sigma_t' = \Sigma_0 + \Sigma_1 X_t^1 + \Sigma_2 X_t^2$, a 4×4 matrix. I use the identification and admissibility constraints given in Joslin

³See Dai and Singleton (2000) for a summary of affine 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.

(2006). 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). I also estimate $A_1(4)$ and $A_2(4)$ models with constraints imposed for unspanned stochastic volatility.

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. (2006a). This specification allows the expected excess returns for exposure to each risk factor to be affine in the state. As elaborated further in Section 4, a flexible market price of risk is critical in matching observed risk premia for holding both bonds and bond options. See Appendix A for the complete model specification.

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

$$E_t^Q[e^{-\int_t^T r_\tau d\tau} f(X_T)]. \quad (2)$$

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

$$P^T(X_t, t) = e^{A(T-t) + B(T-t) \cdot X_t}, \quad (3)$$

where the loadings A and B satisfy the Riccati differential equations

$$\begin{aligned} \dot{B} &= -\rho_1 + K_1^\top B + \frac{1}{2} B^\top H_1 B, & B(0) &= 0, \\ \dot{A} &= -\rho_0 + K_0^\top B + \frac{1}{2} B^\top H_0 B, & A(0) &= 0. \end{aligned} \quad (4)$$

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 coupon bond expiring in P -year, referred to as an in- P -for- Q swaption, may be priced by

$$S_t = E^Q[e^{-\int_t^{t+P} r_\tau d\tau} (CB(X_{t+P}, Q) - 1)^+], \quad (5)$$

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). 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 ODE similar to (4) 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 (1) 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 estimation 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 approximate methods (e.g. Duffie et al. (2003a) and Ait-Sahalia (1999)). However, I estimate the models using full information maximum likelihood estimation. The procedure, developed formally in Joslin and Singleton (2006) uses an extension of the quadrature method used in pricing to implement maximum likelihood. See Appendix C for a summary of the calculations used in the current context.

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 Tables 1 and 2

Table 3 and Table 4 presents the root mean square pricing errors for zeros and swaps. 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. 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.

Table 6 shows the eigenvalues under the risk neutral measure of the drift feedback matrix, $\kappa^Q = -K_1^Q$. 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.

Table 5 gives the pricing errors for swaptions. Data from GovPX indicates that swaption bid-ask spreads range from 1-2% implied volatility. 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% is excluded, the mean square error on

the in 3 months-for 2 year swaption drops to 2.2%. Thus there is in general a good cross sectional fit across the options as well as the yields. The fit for the swaptions maturities and expiries not used in estimation are of similar magnitudes.

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 (μ^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 (μ^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.

However, as shown in Section 4, convexity plays only a small role in bond prices. This means that bond prices depend primarily on risk neutral expectation, μ^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. 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 risks become close to riskless. 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 Role of Convexity in Bond Pricing

Long maturity bonds into terms represent reflecting expectations of future interest rates, risk premia, and convexity effects. The T -year zero coupon yield can be decomposed as

$$y_t^T = y_{t,E} + y_{t,RP} + y_{t,C}, \quad (6)$$

where

$$\begin{aligned} y_{t,E}^T &\equiv \frac{1}{T} \int_t^{t+T} E_t^P[r_\tau] d\tau, \\ y_{t,RP}^T &\equiv \frac{1}{T} \int_t^{t+T} (E_t^Q[r_\tau] - E_t^P[r_\tau]) d\tau, \\ y_{t,C}^T &\equiv -\frac{1}{T} \left(\log E_t[e^{-\int_t^{t+T} r_\tau d\tau}] + \int_t^{t+T} E_t^Q[r_\tau] d\tau \right). \end{aligned} \quad (7)$$

The expectations term represents the bond price discounting with a yield to maturity equal to the average expected future short rate. The convexity term represents the difference in Jensen's inequality from $\exp(E^Q[-\int_0^T r_\tau d\tau]) \neq E^Q[\exp(-\int_0^T r_\tau d\tau)]$.

Figure 1 plots the decomposition of the 2-year, 5-year, and 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 premia effect and that the convexity effects are quite small, even for long maturities.

Table 8 gives the model implied mean convexity effect across models and maturities. Table 9 provides the standard deviation of changes in the convexity effects. For comparison, an $A_1(3)$ model, estimated on the same data but not inverting the swaption, is added to both tables. Tables 8 and 9 again indicate that the convexity effect is nearly negligible for 2-year zeros and small for even 10-year zeros. However, the variation in the convexity effect is very small, even up to 30-year yields. In the $A_1(3)$ model where the swaption is not priced exactly, the variation in the convexity effect is lower still.

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 (4) by eliminating the quadratic convexity term,

$$\dot{B} \approx -\rho_1 + K_1^\top B. \quad (8)$$

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.⁴

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.⁵ 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 10 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 very long maturity bonds, not used in model estimate. Consistent with the rejection of the restriction of two persistent risk factors, Table 3 show that the USV models have very large errors in pricing the long maturities bonds.

As Collin-Dufresne et al. (2006a) argue, the fact that convexity effects are small suggests that volatility may be poorly identified from the cross section of bond prices.⁶ Indeed, even in a model where the constraints for USV are strongly violated, volatility may only be identified through the convexity

⁴More formally, the precise condition is the existence of an eigenvector of K_1^Q which is orthogonal to ρ_1 and loads on the volatility factors.

⁵Joslin (2006) shows that in order for the convexity effect to cancel, three types of restrictions must hold: (1) some factor mean reversion must be 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.

⁶In fact, Andersen and Benzoni (2005) stress the apparent theoretical deficiency of general affine models to produce low correlation between volatility changes and yield changes.

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 multicollinearity) and thus nearly non-singular.

To highlight the mechanism, define

Defintion 4.1. *A diffusive Q -short rate model,*

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

has a weakly spanned volatility risk factor if there exists a change of variable $Y_t = f(X_t)$ such that Y_t^1 drives volatility but does not affects expectations of future interest rates ($\partial\mu/\partial y^1 \equiv 0$)

A non-degenerate affine term structure model of form 1 has a weakly spanned volatility risk factor if there exists an eigenvector of K_1 which is orthogonal to δ_1 . However, not all classes of term structure models admit weakly spanned volatility risk factors. 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.

Note also that the mechanism at work in weakly spanned and unspanned volatility are very different. In USV models, the cross-sectional convexity effects induced by any stochastic variance or covariance effects must be canceled across all maturities. This requires restrictions on the number of stochastic convexity 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. In contrast, the simple conditions of the weakly spanned model require only 1 parameter constraint.

5 Hedging Volatility Risk

As indicated in Section 4, the fact that convexity effects are small means that, even in an unconstrained model, there may be a component of volatility which is largely unrelated to the yield curve. To examine the volatility risk which is generally separate from yield risk, I define the the residual variance

$$V_t^R = V_t - \alpha_1 y_t^{6m} - \alpha_2 y_t^{2y} - \alpha_3 y_t^{10y} \quad (9)$$

where V_t is the variance of the 5-year zero and α is chosen so that V_t^R is locally uncorrelated with the yields $(y^{3m}, y^{2y}, y^{10y})$: $\alpha = \Sigma_y^{-1} \Sigma_{V_y}$. There will be two time scales which will be relevant for the residual volatility. For example, if one purchases a 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. I will refer to the residual volatility here as the 3-month across 9-month residual volatility which will be uncorrelated at the 3-month horizon with the fixed yields. The local residual variance will refer 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.

Figure 3 shows 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 $A_1(4)$ model.⁷ In USV models, this loading will be exactly zero since fixing $N - 1$ yields fixes the entire yield curve. 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

⁷These loadings are found by transforming the original risk factors X_t from the drift-normalize model to the risk factors $Y_t = (y_t^{3m}, y_t^{2y}, y_t^{10y}, V_t^R) = C + DX_t$. The new loadings for maturity τ are transformed by $B(\tau) \mapsto (D^{-1})^\top B(\tau)$

the yield curve. Thus one would anticipate that a fourth factor likely would have only a small effect on yields. Interestingly, the residual variance has a small effect on long maturity yields supporting the analysis in Section 4 that the variation in convexity effects is quite small even at longer maturities.

Figure 4 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 is explaining little of the variation in the yield volatility. The average correlation is 24.6%, indicating that much of the residual variance risk is correlated with the risk in the 6-month, 2-year, and 10-year yields.

Figure 5 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. Also plotted is the fraction of variance in the quoted prices of an at-the-money swaption. This differs from the previous swaption in that the strike is not fixed but rather updated as the yield curve moves. As 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 locally a hedge of a swaption straddle using 6-month, 2-year, and 10-year bonds will be fairly successful.

Figure 6 shows the sensitivity of a 1-year into 5-year swaption as the moneyness is varied. Typical weekly volatility for the 5-year swap rates range from 12 to 18 basis points. 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 yield curve 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 correlation with swaps which have extremely low trading costs.

6 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. Almeida et al. (2005) suggest that bond options identify risk premia. The risk premia for holding a τ -year bond can be computed as

$$\text{risk premium} = B(\tau) \cdot (\mu^P(X_t) - \mu^Q(X_t)) \quad (10)$$

Section 5 indicates that some component of volatility risk is only weakly spanned by bonds. This suggests the possibility that these residual variance risk also drive risk premia for holding long term bonds. To examine this, we can re-express the state variable in terms of $Y_t = (y^{6m}, y^{2y}, y^{10y}, V_t^R)$ and decompose the risk premia into a component associated with the yield curve and a component due to the residual variance. 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 decided to raise 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 risk premia or declining risk premia. Figure 1 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. Figure 2 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. 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 observations support the empirical results that volatility is an important determinant of the risk premium demanded for holding long maturity bonds.

7 Pricing Volatility Risk

Section 6 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, the compensation agents demand for exposure to volatility risk is time-varying and can be inferred from the prices of securities.

Again, considering the state vector $Y_t = (y^{6m}, y^{2y}, y^{10y}, V_t^R)$ we can decompose the risk premium associated with exposure to the V_t^R risk factor. Note that in the $A_1(4)$, the risk premium for exposure to variance risk is necessarily only a function of the variance. This is due to the admissibility constraints required to have a well-defined model. The preferred $A_2(4)$ model shows that primarily variance risk premia depend on the level of volatility and the slope of the yield curve.

Figure 4 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 P -expected value rather than the 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 there variance risk premium depends only on the slope of the yield curve.⁸ 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 6 also suggest that it is difficult to be exposed directly to pure volatility risk over a long horizon and that volatility risk can be partially hedged 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. Fig-

⁸Formally, it is the expectation computed assuming the dynamics follow the measure \hat{P} which differs from the Q -measure by setting the drift of V_t^R equal to its Q drift except for the coefficient on the slope which maintains the coefficient under P

Figure 7 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 .3. This suggests 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.

8 Expected Excess Returns

These results (and related results of Cochrane and Piazzesi (2006), Ludvigson and Ng (2006), Rudebusch et al. (2006), Ludvigson and Ng (2007) and others) suggest that the cross-section of bond yields may poorly identify expected excess returns for holding long maturity bonds. This raises the question of whether expected excess returns may be accurately identified using all fixed income security prices – including fixed income derivatives. To see why it may be the case that expected excess returns are not identified from prices, consider the simple two-factor short rate model where the short rate, r_t , is simply the sum of observable inflation, i_t , and a nominal-real spread, s_t . Suppose that the state variable $X_t = (i_t, s_t)$ follows the diffusion

$$dX_t = \begin{bmatrix} .2 & 0 \\ 0 & .1 \end{bmatrix} (\theta - X_t) dt + \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_s \end{bmatrix} dB_t^P \quad (11)$$

Suppose also that the degree to which agents are risk averse to inflation risk depends on the current level of inflation ($\lambda_t = \lambda_0 + \lambda_1 i_t$) so that under Q :

$$dX_t = \begin{bmatrix} .1 & 0 \\ 0 & .1 \end{bmatrix} (\theta - X_t) dt + \begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_s \end{bmatrix} dB_t^Q \quad (12)$$

In this case, risk-neutrally, i_t and s_t affect fixed income security prices only through their sum and as far as prices are concerned this is a one-factor model. However, knowing both i_t and s_t are informative for learning about expected excess returns for holding long maturity bonds. That is, fixed income security prices (both bonds and derivatives) *do not* determine the state of the economy and *do not* identify expected excess returns.

Indeed, Ludvigson and Ng (2006) 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. In the spirit of Definition 4.1, excess returns will be weakly identified from bond prices when there is a risk factor (after a change of variables) which does not affect Q -expectation of future interest rates but does affect P -expectations of future interest rates. In the example, we change the variable to $Y_t = (r_t, i_t)$. Then fixing r_t and changing i_t do not affect future Q -expectations of interest rates but does affect future P -expectations. This is because the drift of r_t is $\kappa\theta - .2i_t - .1s_t = \kappa\theta - .1i_t - .1r_t$.⁹ Additionally, there is the possibility that excess returns may be identified accurately not through bond prices but through derivative prices. This will be precluded when there another orthogonality condition on the volatility is obtained.

The econometric method used here precludes such identification of risk premia since it is assumed that the state of the economy can be inferred from fixed income security prices. To allow for such flexibility would require the use of a filtering technique which is complicated both by computational burden and by the non-linearity in the option prices.

9 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. However, I find empirical evidence rejecting conditions for unspanned volatility. The residual component of volatility represents a priced risk which also drives the variation in excess returns for holding long maturity bonds. Additionally, I develop computational methods for pricing options and extend the technique to provide maximum likelihood estimation of general affine diffusions.

⁹Precisely, the condition in an affine model is the existence of an eigenvector of κ^Q which is orthogonal to ρ_1 but which is not an eigenvector of κ^P

A Model Specification

For the affine term structure model

$$\begin{aligned} r_t &= \rho_0 + \rho_1 \cdot X_t, \\ dX_t &= \mu_t dt + \sigma_t dB_t, \\ \mu_t &= K_0 + K_1 X_t, \\ \sigma_t \sigma_t' &= H_0 + H_1 \cdot X_t, \end{aligned}$$

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 $\Theta = (\rho_0, \rho_1, K_0^P, K_1^P, K_0^Q, K_1^Q, C_{M+1}^G, C_{M+2}^G, \dots, C_N^G)$, 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):

$$\begin{aligned}
 G(y) &= E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T} \{\delta \cdot X_T \leq y\}], \\
 \hat{G}(y) &= E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T + i\delta \cdot X_T}], \\
 G(y) &= \frac{\hat{G}(0)}{2} - \frac{1}{\pi} \int_0^\infty \frac{1}{t} \text{Im}(\hat{G}(t)e^{ity}) dt.
 \end{aligned} \tag{13}$$

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

$$\frac{dF}{dQ} = \frac{e^{-\int_0^P r_\tau d\tau + d \cdot X_T}}{E^Q[e^{-\int_0^P r_\tau d\tau + d \cdot X_T}]} . \tag{14}$$

More precisely, $\hat{G}(t) \approx ce^{-\sigma_Y^2 t^2/2 + it\mu_Y}$.¹⁰ In the case of an $A_0(N)$ Gaussian model, this equation is exact. Considering this case for now, the Levy integral then becomes:

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

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.

¹⁰It 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:

$$\begin{aligned}
 I &= \int_0^\infty \frac{1}{t} \underbrace{\text{Im}(\hat{f}(t)e^{-ity}e^{\sigma^2 t^2})}_{g(t)} e^{-\sigma^2 t^2} dt \\
 &\approx \sum_i g(t_{i,\sigma}) w_{i,\sigma}
 \end{aligned}$$

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 $b \cdot X_T$.¹¹ 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 11 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

¹¹The 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:

$$\begin{aligned}\hat{f}(s|x_t) &= E[e^{is \cdot X_{t+1}} | X_t = x_t] \\ f(x_{t+1}|x_t) &= \frac{1}{(2\pi)^n} \int \hat{f}(s|x_t) e^{-ix_{t+1} \cdot s} ds\end{aligned}\tag{15}$$

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,

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

If we define $dT/dP = e^{a \cdot X_{t+1}}/E_t[e^{a \cdot X_{t+1}}]$, under T , $E_t[X_{t+1}] \approx \mu + \Sigma a$.¹² 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 $e^{-\frac{1}{2}s^\top \Sigma_t s}$. Thus the integral

$$\begin{aligned}f^T(x_{t+1}) &= \frac{1}{(2\pi)^n} \int \hat{f}^T(s) e^{-is \cdot x_{t+1}} ds \\ &= \frac{1}{(2\pi)^n} \int w(s) e^{-\frac{1}{2}s^\top \Sigma_t s} ds\end{aligned}$$

¹²Note 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.

where $w(s) = e^{-\frac{1}{s^\top \Sigma s} / \hat{f}^T(s)} e^{-is \cdot x_{i+1}} ds \approx 1$. 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).

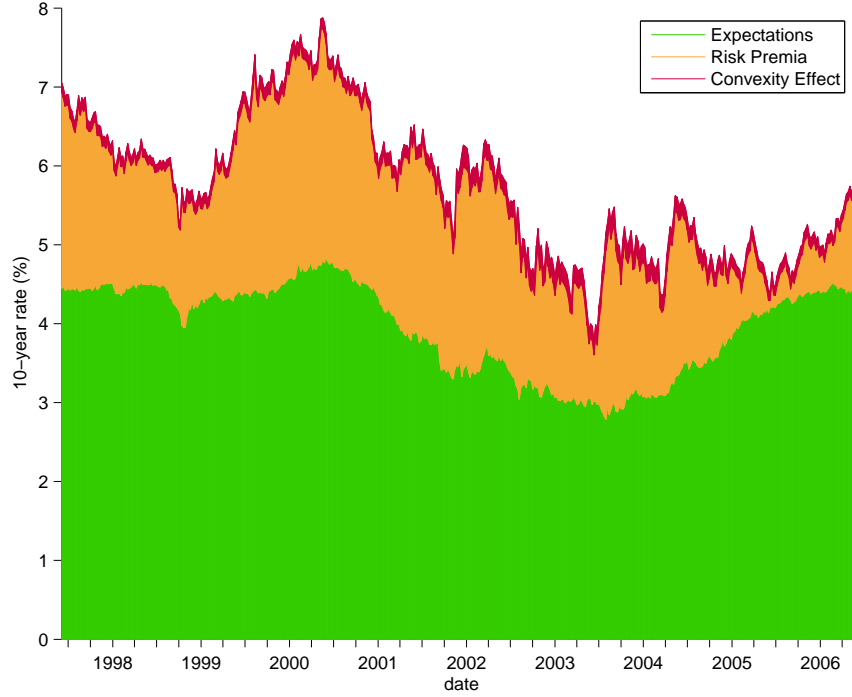


Figure 1: Yield Decomposition

This figure plots model-implied decomposition of the 10-year yield into expectations, risk premium, and convexity components for the estimated $A_1(4)$ model.

$$\begin{aligned}
 y_{\text{expectation}} &= \frac{1}{10} \int_t^{t+10} E_t^P[r_\tau] d\tau, \\
 y_{\text{risk premia}} &= \frac{1}{10} \int_t^{t+10} (E_t^Q[r_\tau] - E_t^P[r_\tau]) d\tau, \\
 y_{t,\text{convexity}} &= -\frac{1}{10} \left(\log E_t[e^{-\int_t^{t+10} r_\tau d\tau}] + \int_t^{t+10} E_t^Q[r_\tau] d\tau \right).
 \end{aligned}$$

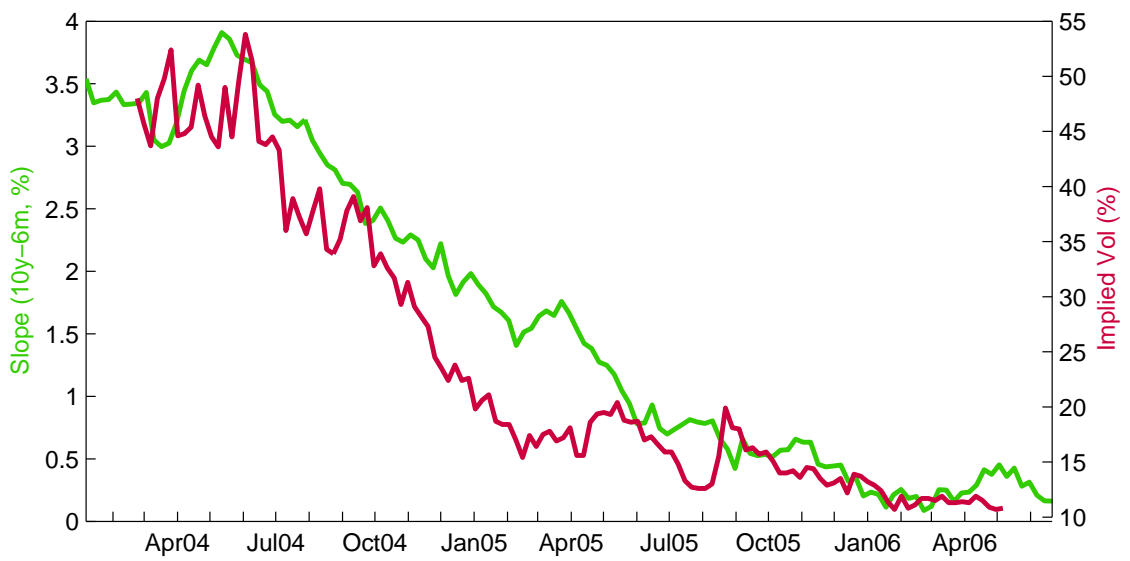


Figure 2: Long-Yield Conundrum

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.

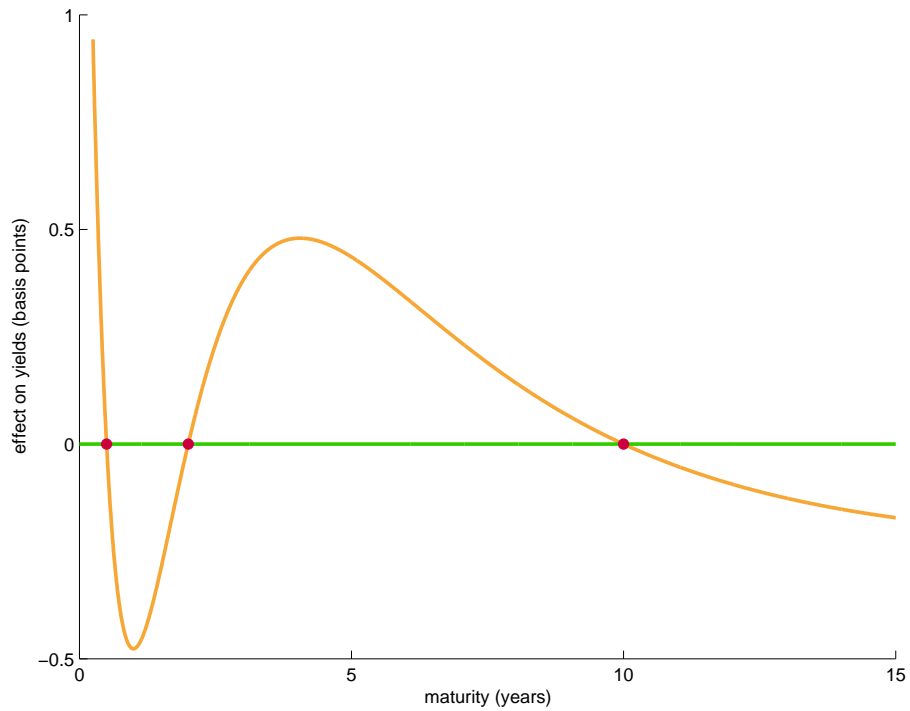


Figure 3: Effect of Residual Variance on the Yield Curve

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.

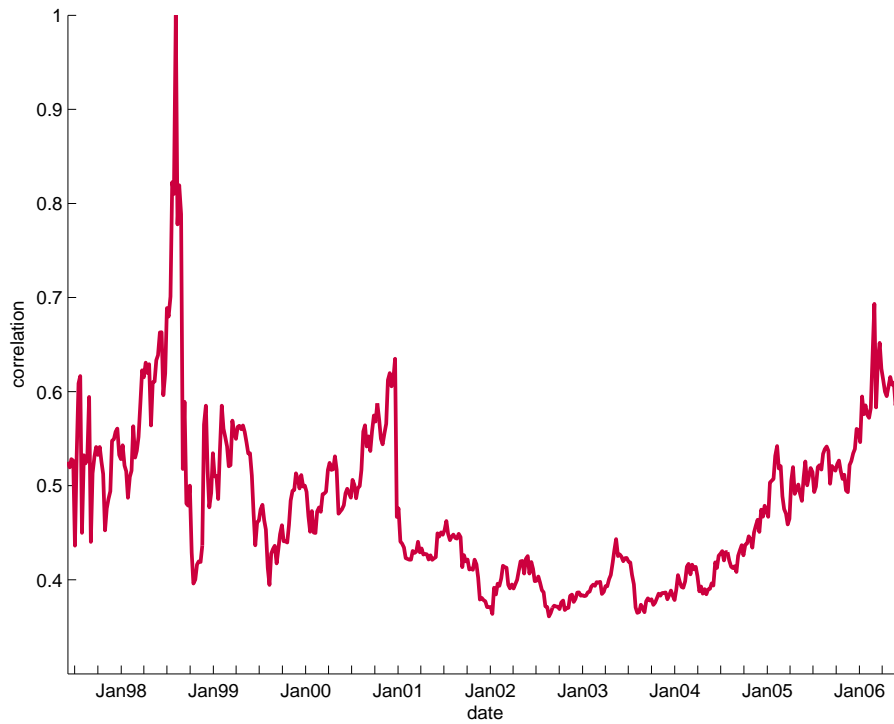


Figure 4: Correlation of Variance with Residual Variance

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 5: 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 locally 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 between 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.

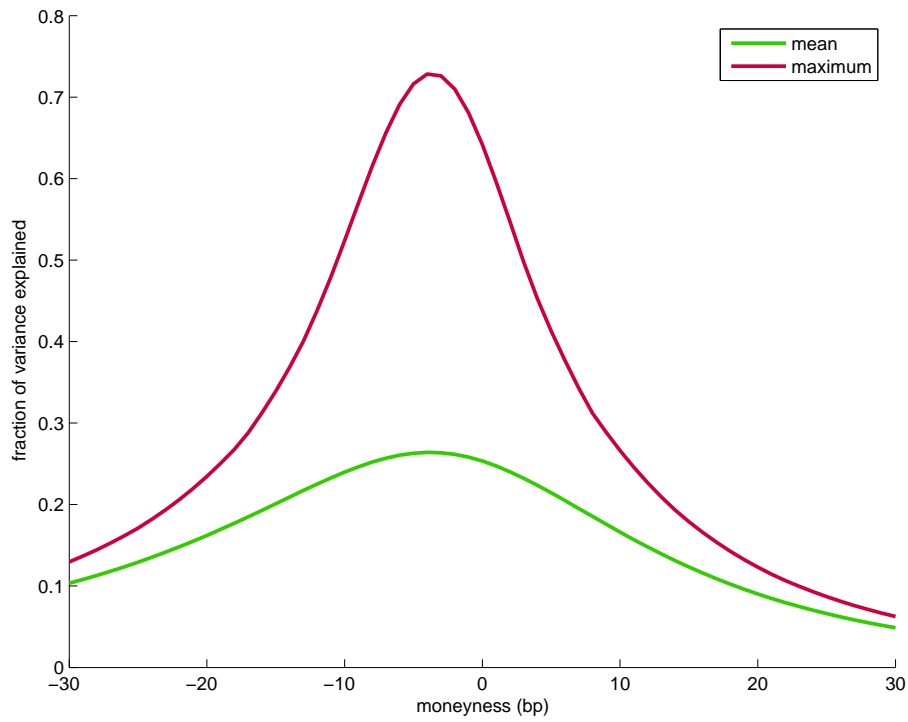


Figure 6: Reduction in Swaption Variance by Yield Correlation

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.



Figure 7: Price of Variance Risk

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

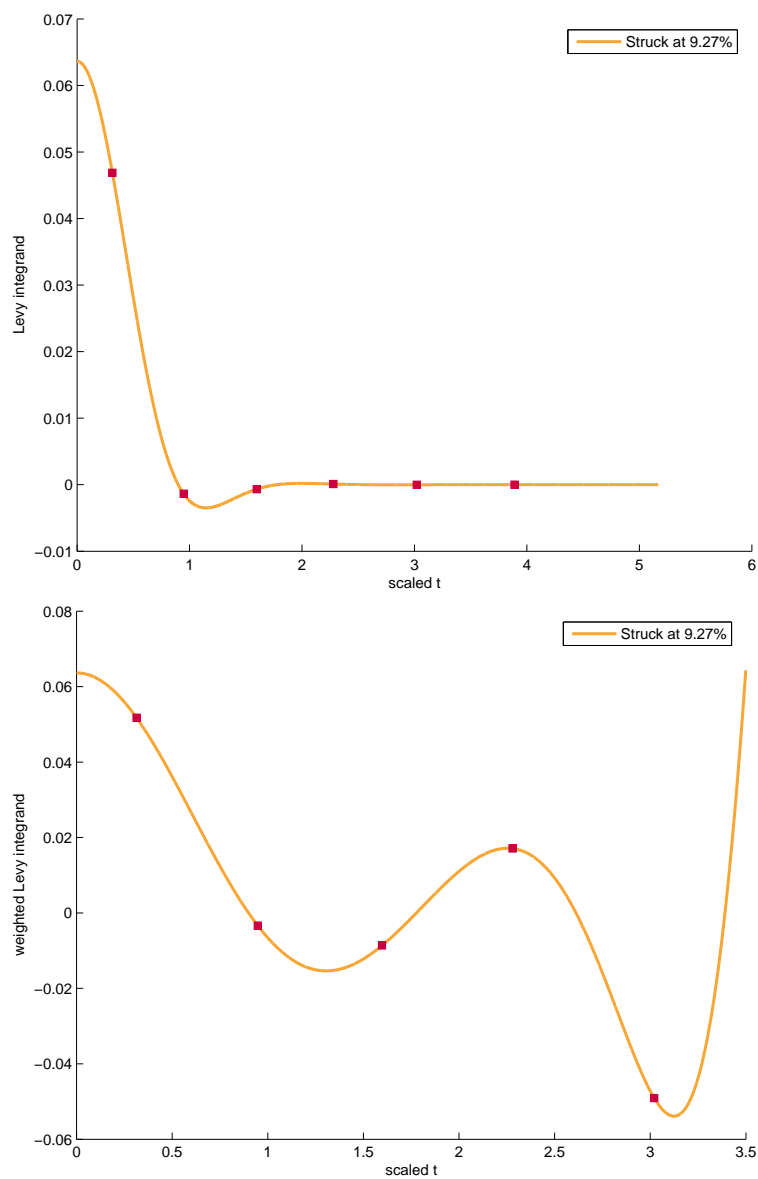


Figure 8: Levy Integrand

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 integrand. The squares indicate nodes used in Gauss-Hermite quadrature.

	$A_1(4)$	$A_2(4)$	$A_1(4)^{USV}$	$A_2(4)^{USV}$
$K_{0,1}^Q$	0.5247	0.5	0.5	1.751
$K_{0,2}^Q$	0	0.5	0	1.054
$K_{0,1}^Q$	0.5247	0.5	0.5	1.751
$K_{0,2}^Q$	0	0.5	0	1.054
$K_{1,1}^Q$	-0.3174	-0.8255	-0.2439	-3.833
$K_{1,2}^Q$	0	0.4036	0	0
$K_{1,3}^Q$	0	0	0	0
$K_{1,4}^Q$	0	0	0	0
$K_{2,1}^Q$	0.08395	0.8716	0.1031	0.648
$K_{2,2}^Q$	-0.6417	-1.085	-0.03375	-1.628
$K_{2,3}^Q$	0	0	-0.6808	0
$K_{2,4}^Q$	0	0	-3.998	0
$K_{3,1}^Q$	0.3874	1.673	-0.02381	-1.217
$K_{3,2}^Q$	0	-1.491	0	0.08898
$K_{3,3}^Q$	-0.05028	-0.6241	-0.0675	-0.1183
$K_{3,4}^Q$	0	0	-0.9504	-0.4021
$K_{4,1}^Q$	0.1344	0.7582	0	-1.069
$K_{4,2}^Q$	0	0.03494	0	-0.03562
$K_{4,3}^Q$	0	0	0	0
$K_{4,4}^Q$	-1.492	-0.05267	-1.233	-0.2366
$K_{1,1}^Q$	-0.3174	-0.8255	-0.2439	-3.833
$K_{1,2}^Q$	0	0.4036	0	0
$K_{1,3}^Q$	0	0	0	0
$K_{1,4}^Q$	0	0	0	0
$K_{2,1}^Q$	0.08395	0.8716	0.1031	0.648
$K_{2,2}^Q$	-0.6417	-1.085	-0.03375	-1.628
$K_{2,3}^Q$	0	0	-0.6808	0
$K_{2,4}^Q$	0	0	-3.998	0
$K_{3,1}^Q$	0.3874	1.673	-0.02381	-1.217
$K_{3,2}^Q$	0	-1.491	0	0.08898
$K_{3,3}^Q$	-0.05028	-0.6241	-0.0675	-0.1183
$K_{3,4}^Q$	0	0	-0.9504	-0.4021
$K_{4,1}^Q$	0.1344	0.7582	0	-1.069
$K_{4,2}^Q$	0	0.03494	0	-0.03562
$K_{4,3}^Q$	0	0	0	0
$K_{4,4}^Q$	-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.

	$A_1(4)$	$A_2(4)$	$A_1(4)^{USV}$	$A_2(4)^{USV}$
$H_{0,1,1}$	0	0	0	0
$H_{0,2,1}$	0	0	0	0
$H_{0,2,2}$	8.951	0	1	0
$H_{0,3,1}$	0	0	0	0
$H_{0,3,2}$	5.518	0	0	0
$H_{0,3,3}$	3.411	0.1914	1	0.01838
$H_{0,4,1}$	0	0	0	0
$H_{0,4,2}$	3.699	0	0	0
$H_{0,4,3}$	2.281	-0.4824	0	0.1468
$H_{0,4,4}$	1.534	1.217	1	2.136
$H_{1,1,1}^1$	1	1	1	1
$H_{1,2,1}^1$	0	0	0	0
$H_{1,2,2}^1$	4.426	0	6.925	0
$H_{1,3,1}^1$	0	0	0	0
$H_{1,3,2}^1$	1.35	0	0	0
$H_{1,3,3}^1$	8.553	0.6444	0	0.02651
$H_{1,4,1}^1$	0	0	0	0
$H_{1,4,2}^1$	2.473	0	0	0
$H_{1,4,3}^1$	-1.401	-2.763	0	0.02399
$H_{1,4,4}^1$	2.029	11.87	0	5.057
$H_{1,1,1}^2$	0	0	0	0
$H_{1,2,1}^2$	0	0	0	0
$H_{1,2,2}^2$	0	1	0	1
$H_{1,3,1}^2$	0	0	0	0
$H_{1,3,2}^2$	0	0	0	0
$H_{1,3,3}^2$	0	1.415	0	2.153
$H_{1,4,1}^2$	0	0	0	0
$H_{1,4,2}^2$	0	0	0	0
$H_{1,4,3}^2$	0	-4.504	0	0
$H_{1,4,4}^2$	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.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
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.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
1 Year	7.5	12.9	7.5	10.5
2 Year	0.1	0.2	0.1	0.2
3 Year	3.9	9.9	3.9	6.0
4 Year	5.0	14.5	5.0	7.9
5 Year	5.1	15.8	5.1	8.0
7 Year	3.7	12.9	3.8	6.1
10 Year	0.7	2.4	0.7	1.3
12 Year	2.6	7.9	2.6	4.4
15 Year	6.2	26.9	6.3	13.4
20 Year	8.9	67.4	9.1	25.7
25 Year	9.0	124.1	9.3	34.3
30 Year	10.4	205.6	10.6	41.7

Table 4: Swap Pricing Errors

Root mean square swap rate pricing errors in basis points. 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 5: 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.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
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.

	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
1st Eigenvalue	1.86	2.33	4.98*	0.81
2nd Eigenvalue	1.11	0.48*	0.89	0.81
3rd Eigenvalue	0.48	0.31	0.46	0.92*
4th Eigenvalue	0.47*	0.10	0.35*	1.27*

Table 7: Eigenvalues Under The Physical Measure

Eigenvalues of the mean reversion matrix, κ^P , under the physical 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.

	$A_1(3)$	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
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

Table 8: Average Convexity Effects

This table gives the sample mean model-implied convexity effects, in basis points, for different bond maturities. The convexity effect of an T -year yield:

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

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.

	$A_1(3)$	$A_1(4)$	$A_1(4)^{USV}$	$A_2(4)$	$A_2(4)^{USV}$
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

Table 9: Time Variation of Convexity Effects

This table gives the sample standard deviation, in basis points, of weekly changes in the model-implied convexity effects for different bond maturities. 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.

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

Table 10: 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.

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 11: 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.

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.
- Caio Almeida, Jeremy Graveline, and Scott Joslin. Do options contain information about excess bond returns? *Working Paper, Stanford University*, 2005. Working Paper, Stanford University.
- Torben Andersen and Luca Benzoni. Can bonds hedge volatility risk in the u.s. treasury market? a specification test for affine term structure models. Working Paper, December 2005.
- Ruslan Bikbov and Mikhail Chernov. Term structure and volatility: Lessons from the eurodollar markets. Working Paper, December 2005.
- 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 a ne models: Theory and evidence. *Journal of Financial Economics*, 2006a. Forthcoming.

- Patrick Cheridito, Damir Filipovic, and Robert L. Kimmel. A note on the dai-singleton canonical representation of affine term structure models. Working Paper, Princeton University, October 2006b.
- John Cochrane and Monika Piazzesi. Decomposing the yield curve. Working Paper, November 2006.
- 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. Working Paper, 2006a.
- Pierre Collin-Dufresne, Robert S. Goldstein, and Christopher S. Jones. Identification of maximal affine term structure models. Working Paper, August 2006b.
- 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.
- Jefferson Duarte. Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies*, 17(2):379–404, 2003.
- Gregory R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):405–443, February 2002.

- 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.
- Rong Fan, Anurag Gupta, and Peter Ritchken. On pricing and hedging in the swaption market: How many factors, really? Working Paper, 1 October 2001.
- Massoud Heidari and Liuren Wu. Are interest rate derivatives spanned by the term structure of interest rates? *Journal of Fixed Income*, 13(1):75–86, June 2003.
- 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.
- Scott Joslin. Can unspanned stochastic volatility models explain the cross section of bond volatilities? *Working Paper, Stanford University*, 2006. Working Paper, Stanford University.
- Scott Joslin and Kenneth J. Singleton. Work in progress. 2006. Working Paper.
- 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. Working Paper, August 2006.
- Sydney C. Ludvigson and Serena Ng. Bond risk premia and realized jump volatility. Working Paper, February 2007.
- 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.
- Samuel Thompson. Identifying term structure volatility from the libor-swap curve. Working Paper, December 2004.