

Do Interest Rate Options Contain Information About Excess Returns?

Caio Almeida^a, Jeremy J. Graveline^b, Scott Joslin^c

^a*Graduate School of Economics, Fundação Getulio Vargas*

^b*University of Minnesota, Carlson School of Management*

^c*MIT Sloan School of Management*

Abstract

There is strong empirical evidence that long-term interest rates contain a time-varying risk premium. Options may contain valuable information about this risk premium because their prices are sensitive to the underlying interest rates. We use the joint time series of swap rates and interest rate option prices to estimate dynamic term structure models. The risk premiums that we estimate using option prices are better able to predict excess returns for long-term swaps over short-term swaps. Moreover, in contrast to previous literature, the most successful models for predicting excess returns have risk factors with stochastic volatility. We also show that the stochastic volatility models we estimate using option prices match the failure of the expectations hypothesis.

Keywords: interest rates, options, risk premia, excess returns, forecasting
JEL Codes: C53, C58, E43, E47, G12

Email addresses: calmeida@fgv.br (Caio Almeida), jeremy@umn.edu (Jeremy J. Graveline), sjoslin@mit.edu (Scott Joslin)

Preprint submitted to Journal of Econometrics

October 18, 2010

1. Introduction

A bond or swap that is sold before it matures has an uncertain return. There is strong empirical evidence that this return is predictable and time-varying which suggests that long-term interest rates contain a time-varying risk premium.¹ In this paper, we ask whether interest rate option prices can be used to obtain better estimates of this risk premium. Options may contain valuable information because their prices are sensitive to the volatility and risk premiums in the underlying interest rates.

We use an arbitrage-free term structure model for our empirical analysis because it describes the joint dynamics of interest rate option prices and the underlying interest rates. Related papers that use term structure models to investigate the risk premium in long term interest rates use only bonds or swap rates for estimation.² Instead, we use the joint time series of both swap rates and interest rate cap prices with different maturities. Our main finding is that the risk premiums we estimate using option prices are better able to predict excess returns for long-term swaps. To measure predictability, we compute an R^2 -statistic that compares the squared differences between expected returns and 3-month realized returns on zero-coupon swaps with different maturities. However, rather than compute expected returns using linear projection, we instead use our model-implied expected returns as a more stringent metric. Across different maturities, this measure of predictability typically doubles when we use options to estimate a term structure model with one stochastic volatility factor. The improvement is almost threefold for a term structure model with two stochastic volatility factors.

Risk premia in term structure models reflect the market prices of risk for the factors driving interest rates (or equivalently, their risk-neutral dynamics). To better understand the improvement in predictability, we examine the estimated risk premia and find that the risk-neutral dynamics of the factors are quite different when the models are estimated with and without options. In the models that are estimated without options, long-run interest rates under the risk-neutral measure are high (10.4–13%) and have a very

¹See [Fama and Bliss \(1987\)](#), [Campbell and Shiller \(1991\)](#), and [Cochrane and Piazzesi \(2005\)](#).

²See [Duffee \(2002\)](#), [Dai and Singleton \(2002\)](#), [Duarte \(2004\)](#), and [Cheridito et al. \(2007\)](#).

slow rate of mean reversion with a half-life of 22.4 to 31.4 years. In contrast, the models estimated with options have a more modest risk-neutral long-run mean of 8% with a rate of mean reversion on the order of business cycle frequency (the half-life is about 8 years). Both of these combinations have similar implications for swap rates, which makes them difficult to distinguish using only swaps. However, the two mechanisms are clearly differentiated by their implications for interest rate option prices. Therefore, option prices allow for a more precise identification of the market prices of risk and the models that are estimated with options are better able to explain variation in excess returns.

Option prices also help to resolve the tension, identified in previous papers, between matching the first and second moments of bond returns. Both [Duffee \(2002\)](#) and [Cheridito et al. \(2007\)](#) find that excess bond returns are best captured by constant volatility models that cannot match the time-series variation in interest rate volatility. We study the same 3-factor affine term structure models that were developed by [Cheridito et al. \(2007\)](#) (a generalization of the models developed and studied by [Duffee, 2002](#)), but we use the joint time series of swaps and interest rate cap prices to estimate the models.³ When we use option prices to estimate the model parameters, the models with one or two stochastic volatility factors better capture the variation in interest rate volatility *and* are also best at predicting excess returns.

The models with stochastic volatility that we estimate with options also satisfy two additional challenges posed by [Dai and Singleton \(2003\)](#): they successfully price interest rate caps and they capture the failure of the expectations hypothesis.⁴ [Dai and Singleton \(2002\)](#) find that only models with constant volatility successfully match the failure of the expectations hypothesis. We show that term structure models with stochastic volatility also match

³Swaps are based on the same LIBOR interest rates as caps, which is why we choose to model swap rates rather than government bond yields. [Dai and Singleton \(2000\)](#) also note that the institutional features that affect government bond yields are not accounted for in standard term structure models.

⁴The failure of the expectations hypothesis refers to the empirical property that excess returns to long-term bonds and swaps are negatively related to the slope of the yield curve (and increasingly so for longer maturity yields). If the expectations hypothesis holds, which would be the case if investors are risk-neutral with respect to interest rate risk, then forward rates are the best linear predictor of future interest rates. See [Fama \(1984a\)](#), [Fama \(1984b\)](#), [Fama and Bliss \(1987\)](#) and [Campbell and Shiller \(1991\)](#).

the failure of the expectations hypothesis when we include options in estimation. We further analyze these results and show that option prices help to identify the portions of the risk premium that are related to the slope of the yield curve.

Our empirical analysis deviates from recent research that has focused on unspanned stochastic volatility, or USV, in fixed income markets.⁵ In term structure models that exhibit USV, interest rate options have both an econometric and an economic role because they cannot be replicated using the underlying bonds or swaps. We do not consider unspanned stochastic volatility models in this paper because our objective is to focus exclusively on the econometric benefits of using options to estimate the risk premium in long-term interest rates.

Other papers have used options to estimate term structure models, but do not examine their impact on a model's ability to capture the dynamics of interest rates and predict excess returns. [Umantsev \(2002\)](#) estimates affine models jointly using both swaps and swaptions and analyzes the volatility structure of these markets. [Longstaff et al. \(2001\)](#) and [Han \(2007\)](#) explore the correlation structure in yields that is required to simultaneously price both caps and swaptions. [Bikbov and Chernov \(2010\)](#) use both Eurodollar futures and short-dated option prices to estimate affine term structure models and discriminate between various volatility specifications.

Our paper is also related to empirical papers that examine the joint time series of option prices and returns in equity or foreign exchange markets. [Chernov and Ghysels \(2000\)](#), [Pan \(2002\)](#), [Jones \(2003\)](#), and [Eraker \(2004\)](#) analyze S&P 100 or 500 index returns jointly with options on the index. [Bakshi et al. \(2008\)](#) and [Graveline \(2008\)](#) study foreign exchange options and the underlying currency returns. These papers use option prices to help estimate risk premia in the underlying equity or currency returns and our paper has a similar objective applied to fixed income markets.⁶

⁵See [Collin-Dufresne and Goldstein \(2002a\)](#), [Collin-Dufresne et al. \(2009\)](#), [Andersen and Benzoni \(2008\)](#), [Li and Zhao \(2006\)](#), [Thompson \(2008\)](#), [Bikbov and Chernov \(2010\)](#), [Joslin \(2007\)](#), and [Kim \(2007\)](#).

⁶See also [Jackwerth \(2000\)](#), [Aït-Sahalia and Lo \(2000\)](#), and [Aït-Sahalia et al. \(2001\)](#) for papers that compare the risk-neutral distribution of returns implied from option prices to the objective distribution of returns inferred from time-series data.

The remainder of the paper is organized as follows. Section 2 describes the data and estimation procedure we use. Section 3 describes the fit of the models to swap rates, cap prices, and the conditional volatility of interest rates. Section 4 compares how well the models we estimate predict excess returns and Section 5 examines the dependence of expected excess returns on the level, slope, and curvature of the yield curve. Section 6 concludes.

2. Model and Estimation

Our objective in this paper is to examine how well different arbitrage-free affine term structure models predict excess returns for long-term swaps. We consider dynamic term structure models in which the short interest rate, r , is driven by a 3-dimensional latent factor X_t that follows an affine⁷ diffusion process. Following the notation in Duffie (2001), the models can be expressed as

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

$$dX_t = [\mathcal{K}_0^{\mathbb{P}} + \mathcal{K}_1^{\mathbb{P}} X_t] dt + \sigma(X_t) dW_t^{\mathbb{P}}, \quad (1b)$$

$$dX_t = [\mathcal{K}_0^{\mathbb{Q}} + \mathcal{K}_1^{\mathbb{Q}} X_t] dt + \sigma(X_t) dW_t^{\mathbb{Q}}, \quad (1c)$$

where $W_t^{\mathbb{P}}$ ($W_t^{\mathbb{Q}}$) is a 3-dimensional Brownian motion under the historical (risk-neutral) measure and $\sigma(X_t)\sigma(X_t)^\top$ is a 3×3 matrix $H_0 + \sum H_1^k X_t^k$ whose (i, j) entry is given by $H_{0ij} + \sum_{k=1}^3 H_{1ij}^k X_t^k$. The short interest rate and risk-neutral dynamics allow us to price a variety of fixed income instruments. For example, a security that pays $g(X_T)$ at time T will have price at time t given by

$$P_t^g = E_t^{\mathbb{Q}}[e^{-\int_t^T r_s ds} g(X_T)]. \quad (2)$$

Dai and Singleton (2000) discuss admissibility and identification issues for the model specification in equation (1). First, admissibility refers to the fact that, in order to ensure a well-defined process, parameter constraints must be imposed so that the covariance remains positive semi-definite. Second,

⁷Researchers have also extensively studied quadratic term structure models (see Ahn et al., 2002, and Leippold and Wu, 2002). However, Cheng and Scaillet (2007) show that affine and quadratic term structure models are equivalent and therefore our choice to restrict the analysis to affine models is without loss of generality.

due to the latent nature of the states, the parameters may not be econometrically identified, as two distinct sets of parameters can give rise to observationally equivalent models. Dai and Singleton (2000) partition N -factors models into subsets, denoted $A_M(N)$, where M is the number of factors that drive stochastic volatility. We estimate 3-factor term structure models with $M = 0, 1$ or 2 factors driving stochastic volatility. To ensure admissibility we impose the following constraints on the parameters in estimation:

$$\mathcal{K}_{0,i}^{\mathbb{P}} \geq 0, \quad i \leq M; \quad (3a)$$

$$\mathcal{K}_{0,i}^{\mathbb{Q}} \geq 0, \quad i \leq M; \quad (3b)$$

$$\mathcal{K}_{1,ij}^{\mathbb{P}} \geq 0, \quad i, j \leq M, i \neq j; \quad (3c)$$

$$\mathcal{K}_{1,ij}^{\mathbb{Q}} \geq 0, \quad i, j \leq M, i \neq j; \quad (3d)$$

$$\mathcal{K}_{1,ij}^{\mathbb{P}} = 0, \quad i \leq M < j; \quad (3e)$$

$$\mathcal{K}_{1,ij}^{\mathbb{Q}} = 0, \quad i \leq M < j; \quad (3f)$$

$$H_0 \text{ is positive semi-definite}; \quad (3g)$$

$$H_{0,ii} = 0, \quad i \leq M; \quad (3h)$$

$$H_1^k \text{ is positive semi-definite with } H_1^k = 0 \text{ for } k > M; \quad (3i)$$

$$H_{1,ii}^k = 0 \text{ when } i \leq M \text{ and } i \neq k. \quad (3j)$$

Under these constraints, the first M factors are CIR (Cox et al., 1985) processes that drive both volatility (through H_1) and interest rates (through ρ_1), while the remaining $N - M$ factors are conditionally Gaussian with the local conditional volatility determined by the first M factors. Conditions (3a–3f) ensure that the conditional mean of the CIR factors stays positive under \mathbb{P} and \mathbb{Q} : positive constant in the drift, positive feedback from one CIR factor to another, and no feedback from the Gaussian factors (which can be either positive or negative) to the CIR factors. Conditions (3g–3j) ensure that the conditional covariance remains positive semi-definite: the matrices that determine volatility are positive semi-definite, and the volatility of a CIR factor is zero when the level of the factor is zero (which requires both (3i) and (3j)). Notice that conditions (3g) and (3h) together imply that $H_{0,ij} = 0$ for $i \leq M$ and any j , since if the variance of a random variable is zero then so is its covariance with any other random variable.

We also impose the Feller condition so that the factors driving stochastic

volatility remain strictly positive under both \mathbb{P} and \mathbb{Q} :

$$\mathcal{K}_{0,i}^{\mathbb{P}} \geq \frac{1}{2} H_{1ii}^i, \quad i \leq M; \quad (4a)$$

$$\mathcal{K}_{0,i}^{\mathbb{Q}} \geq \frac{1}{2} H_{1ii}^i, \quad i \leq M. \quad (4b)$$

The Feller condition allows the historical and risk-neutral measures, \mathbb{P} and \mathbb{Q} , to be equivalent measures even when $\mathcal{K}_1^{\mathbb{Q}}$ and $\mathcal{K}_1^{\mathbb{P}}$ differ in the upper $M \times M$ block. This allows for a fully flexible market price of risk as in [Cheridito et al. \(2007\)](#) (see also [Liptser et al., 2000](#)).⁸

We impose the following additional constraints (as in [Dai and Singleton, 2000](#)) to obtain econometric identification of the parameters:

$$\rho_{1,i} \geq 0, \quad i > M; \quad (5a)$$

$$H_{0,ij} = 1, \quad i = j > M, \quad 0 \text{ otherwise}; \quad (5b)$$

$$H_{1,kk}^k = 1, \quad k \leq M; \quad (5c)$$

$$H_{1,ij}^k = 0, \quad k \leq M, \quad i \neq j; \quad (5d)$$

$$\mathcal{K}_{0,i}^{\mathbb{P}} = 0, \quad i > M; \quad (5e)$$

$$\mathcal{K}_{1,ij}^{\mathbb{P}} = 0, \quad \text{if } M = 0 \text{ and } j > i. \quad (5f)$$

Condition (5a) fixes the sign of the locally Gaussian factors. Conditions (5b–5d) scale the latent factors and orthogonalize their innovations. Condition (5e) removes a level indeterminacy by translating the Gaussian variables so that their mean under \mathbb{P} is zero. Finally, condition (5f) normalizes the Gaussian variables by using an orthogonal transformation of the variables (which maintains the orthogonal innovations) to generate a lower-triangular feedback for the Gaussian variables, as in the Schur decomposition of a matrix.

[Duffie and Kan \(1996\)](#) show that zero-coupon bond prices are exponential affine in the factors. Specifically, the price at time t of a zero-coupon bond that pays \$1 at time T is given by

$$P_t^T = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \right] = e^{A(T-t) + B(T-t) \cdot X_t}, \quad (6)$$

⁸For all the specifications that we estimate, these inequality constraints do not in fact bind.

where $B(\cdot)$ and $A(\cdot)$ solve the Riccati ODEs

$$\frac{d}{d\tau}B(\tau) = -\rho_1 + (\mathcal{K}_1^{\mathbb{Q}})^\top B(\tau) + \frac{1}{2}B(\tau)^\top H_1 B(\tau), \quad B(0) = 0, \quad (7a)$$

$$\frac{d}{d\tau}A(\tau) = -\rho_0 + (\mathcal{K}_0^{\mathbb{Q}})^\top B(\tau) + \frac{1}{2}B(\tau)^\top H_0 B(\tau), \quad A(0) = 0, \quad (7b)$$

and $B(\tau)^\top H_1 B(\tau)$ is a vector whose k th entry is given by $B(\tau)^\top H_1^k B(\tau)$.

Previous papers that investigate the risk premium in long-term interest rates use only the time series of bonds or swaps with different maturities for estimation. Instead, we also use the time series of interest rate cap prices with different maturities. An interest rate cap is a portfolio of options on the 3-month Libor rate that effectively caps the interest rate paid on the floating side of a swap.⁹ We include cap prices in estimation because they may contain additional econometric information about the risk premium in long-term interest rates. Since caps are interest rate options, their prices are sensitive to the volatility of the risk factors. Cap prices also depend on interest rate risk premia, which are embedded in the risk-neutral distribution of the factors.

The price of an N -period cap with strike rate \bar{C} on 3-month floating interest payments is

$$C_t^N(\bar{C}) = \sum_{n=2}^N \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^{t+0.25n} r_s ds} \underbrace{0.25 (L_{t+0.25(n-1)} - \bar{C})^+}_{\text{caplet payoff}} \right], \quad (8)$$

where $L_{t+0.25(n-1)}$ is the 3-month Libor interest rate so that

$$1 + 0.25L_{t+0.25(n-1)} = 1/P_{t+0.25(n-1)}^{t+0.25n}. \quad (9)$$

Duffie et al. (2000) show that cap prices in affine term structure models can be computed as a sum of inverted Fourier transforms. However, when the solutions A and B to the Riccati ODEs in equation (7) are not known in closed form, direct Fourier inversion can be too computationally expensive

⁹Other papers, such as Umantsev (2002), have used swaptions, which are options on swaps. We choose to use interest rate caps because it is easier to compute their prices without resorting to approximations.

for use in estimation. Instead, we use a more efficient adaptive quadrature method that is based on [Joslin \(2007\)](#).

Our data, obtained from Datastream, consists of weekly Libor rates, swap rates, and at-the-money cap implied volatilities from January 1995 to February 2006. We use 3- and 6-month Libor and the entire term structure of swap rates to bootstrap zero-coupon swap rates at 1-, 2-, 3-, 4-, 5-, 7-, and 10-years.¹⁰ We also use at-the-money caps with maturities of 1-, 2-, 3-, 4-, 5-, 7-, and 10-years.

We use quasi-maximum likelihood to estimate model parameters for $A_0(3)$, $A_1(3)$, and $A_2(3)$ models. Following [Chen and Scott \(1993\)](#), we estimate all of the models under the assumption that 3-month Libor and the 2- and 10-year zero-coupon swap rates are priced exactly, while the remaining rates are priced with error.¹¹ In addition, we estimate another set of parameters for the $A_1(3)$ and $A_2(3)$ models under the assumption that at-the-money caps with maturities of 1-, 2-, 3-, 4-, 5-, 7-, and 10-years are also priced with error. We refer to these versions of the models that we estimate with option prices as the $A_1(3)^o$ and $A_2(3)^o$ models.

Given the dynamics in equation (1b), the conditional mean, $\bar{X}_{t,\Delta t} = \mathbb{E}_t[X_{t+\Delta t}]$, satisfies the differential equation,

$$\partial \bar{X}_{t,u} / \partial u = \mathcal{K}_0^{\mathbb{P}} + \mathcal{K}_1^{\mathbb{P}} \bar{X}_{t,u}, \quad (10a)$$

with the initial condition $\bar{X}_{t,0} = \mathbb{E}_t[X_t] = X_t$. Similarly, the conditional covariance, $V_{t,\Delta t} = \mathbb{E}_t \left[(X_{t+\Delta t} - \bar{X}_{t,\Delta t})(X_{t+\Delta t} - \bar{X}_{t,\Delta t})^\top \right]$, satisfies the differential equation,

$$\partial V_{t,u} / \partial u = \mathcal{K}_1^{\mathbb{P}} V_{t,u} + V_{t,u} \mathcal{K}_1^{\mathbb{P}\top} + H_0 + H_1 \cdot \bar{X}_{t,u}, \quad (10b)$$

with initial condition $V_{t,0} = 0$. This coupled system of linear constant coefficient ordinary differential equations can be solved in closed form.

¹⁰Our bootstrap procedure uses the common assumption that forward swap zero rates are constant between observations.

¹¹By assuming that a subset of securities is priced correctly by the model, we can use these prices to invert for the values of the latent states. [Duffee \(2002\)](#), [Dai and Singleton \(2002\)](#), and [Cheridito et al. \(2007\)](#) also use this approach to invert for the latent states. See [Chen and Scott \(1993\)](#) for more details. We choose the 3-month rate as it is the reference rate for the underlying caps. Moreover, 2-year and 10-year swaps are among the most liquid maturities and capture both the moderate and long end of the term structure.

For quasi-maximum likelihood we assume that, ignoring constants, the log likelihood of the state vector is

$$\mathcal{L}_X = -\frac{1}{2} \sum \left\{ \ln |V_{t,\Delta t}| + (X_{t+\Delta t} - \bar{X}_{t,\Delta t})^\top V_{t,\Delta t}^{-1} (X_{t+\Delta t} - \bar{X}_{t,\Delta t}) \right\}. \quad (11)$$

Let Y_t denote the vector of observed 3-month, 2-year, and 10-year zero-coupon swap rates that we use to invert for the latent states X_t .¹² From equation (6) we can compute $\mathcal{A} \in \mathbb{R}^3$ and $\mathcal{B} \in \mathbb{R}^{3 \times 3}$ (given the parameters) and write

$$Y_t = \mathcal{A} + \mathcal{B}X_t \quad \Rightarrow \quad X_t = \mathcal{B}^{-1} [Y_t - \mathcal{A}]. \quad (12)$$

Using our quasi-maximum likelihood assumption, the log-likelihood of Y_t is

$$\mathcal{L}_Y = \mathcal{L}_X - \sum \ln |\mathcal{B}|. \quad (13)$$

Using equation (6) and the states we have recovered from equation (12), we can compute the model-implied zero-coupon yield for any maturity. Let $\tilde{\varepsilon}_k$ denote the \mathcal{T} -dimensional full time series of the pricing errors (as measured by the difference between the observed yield and the model-implied yield) for the k th-maturity interest rate that is priced with error (6-month Libor and 1-, 3-, 4-, 5-, and 7-year zero-coupon swap rates). We assume that these zero-coupon swap rate pricing errors are independent with mean zero Gaussian distribution so that the log-likelihood, ignoring constants, is

$$\mathcal{L}_{\tilde{Y}} = -\frac{1}{2} \sum_{k=1}^6 \left\{ \mathcal{T} \cdot \ln \tilde{\sigma}_k^2 + \frac{\tilde{\varepsilon}_k^\top \tilde{\varepsilon}_k}{\tilde{\sigma}_k^2} \right\}. \quad (14)$$

Here, we treat the $\tilde{\sigma}_k$ as unknown parameters of the model.

Similarly, using equation (8), let $\tilde{\eta}_k$ denote the \mathcal{T} -dimensional full time series of pricing errors for the k th-maturity interest rate cap (we price caps with maturities 1-, 2-, 3-, 4-, 5-, 7-, and 10-years). Again, we assume that these cap pricing errors are independent with mean zero Gaussian distribution so that the log-likelihood, ignoring constants, is

$$\mathcal{L}_{\tilde{C}} = -\frac{1}{2} \sum_{k=1}^7 \left\{ \mathcal{T} \cdot \ln \tilde{v}_k^2 + \frac{\tilde{\eta}_k^\top \tilde{\eta}_k}{\tilde{v}_k^2} \right\}. \quad (15)$$

¹²Our exposition of the estimation procedure draws from [Fisher and Gilles \(1996\)](#).

We treat the \tilde{v}_k as unknown parameters of the model.

Finally, we choose the parameters to maximize the log-likelihood of the yields that are priced exactly, the yields that are priced with errors, and the caps that are priced with error,

$$\mathcal{L} = \mathcal{L}_Y + \mathcal{L}_{\tilde{Y}} + \mathcal{L}_{\tilde{C}}. \quad (16)$$

Our estimates of the parameters are provided in Tables 1 and 2. Standard errors are computed by the BHHH method using the outer product of the gradient (see Davidson and MacKinnon, 1993). It is difficult to directly assign meaning to many of the parameters since they govern the dynamics of a latent process. That is to say, we are more interested in functions of the parameters that pertain to quantities of economic interest. To this end, we now proceed in Section 3 to compare and contrast the implications of the estimated models for the cross-sectional properties of the yield curve and interest rate option prices. We also examine the implications for the conditional volatility of yields. In Sections 4 and 5, our main focus turns to the predictability of bond returns across the estimated models.

3. Fit to Prices and Conditional Volatility

In this section we examine how well the term structure models match zero-coupon swap rates, cap prices, and the conditional volatility of interest rates.

Table 3 provides the root mean squared pricing errors (in basis points) for zero-coupon swap rates with different maturities. The root mean squared errors are 0 for the 3-month, 2-, and 10-year zero-coupon swap rates because the latent state variables are chosen so that the models correctly price these rates. The root mean squared pricing errors for other maturities range from about 4 basis points to about 10 basis points, with slightly larger errors for the short maturities and a better fit for longer maturities.¹³ There is very little difference in the cross-sectional fit between the $A_0(3)$, $A_1(3)$, and

¹³The cross-sectional pricing errors for all of the models that we estimate are comparable with the pricing errors reported in recent papers such as Dai and Singleton (2000), Duffee (2002), and Cheridito et al. (2007).

Parameter	Model				
	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
$\mathcal{K}_{0,1}^{\mathbb{P}}$	0	3.61 (1.5)	2.26 (1.7)	0.571 (2.4)	0.799 (3.6)
$\mathcal{K}_{0,2}^{\mathbb{P}}$	0	0	0	1.32 (4.2)	0.944 (2.3)
$\mathcal{K}_{0,3}^{\mathbb{P}}$	0	0	0	0	0
$\mathcal{K}_{0,1}^{\mathbb{Q}}$	1.39 (25)	1.10 (0.14)	1.30 (0.31)	1.92 (1.2)	1.38 (0.15)
$\mathcal{K}_{0,2}^{\mathbb{Q}}$	0.402 (9.4)	0.825 (0.44)	4.26 (1.2)	0.899 (0.86)	0.721 (0.52)
$\mathcal{K}_{0,3}^{\mathbb{Q}}$	-0.227 (2.2)	-0.254 (5.8)	2.54 (1.0)	-0.300 (0.63)	0.315 (1.2)
$\mathcal{K}_{1,11}^{\mathbb{P}}$	-0.0277 (0.28)	-1.53 (0.62)	-5.09e-5 (0.49)	-0.757 (0.44)	-0.991 (1.2)
$\mathcal{K}_{1,12}^{\mathbb{P}}$	0	0	0	1.03 (0.57)	0.495 (2.1)
$\mathcal{K}_{1,13}^{\mathbb{P}}$	0	0	0	0	0
$\mathcal{K}_{1,21}^{\mathbb{P}}$	0.664 (0.41)	0.447 (0.59)	-0.682 (0.74)	0.493 (0.85)	0.585 (0.89)
$\mathcal{K}_{1,22}^{\mathbb{P}}$	-0.340 (0.56)	-0.592 (0.67)	-1.05 (0.37)	-1.03 (1.1)	-1.45 (1.0)
$\mathcal{K}_{1,23}^{\mathbb{P}}$	0	6.78e-3 (0.53)	-1.76 (1.0)	0	0
$\mathcal{K}_{1,31}^{\mathbb{P}}$	-0.947 (0.49)	-0.886 (1.1)	-0.486 (0.42)	-0.245 (0.092)	-0.699 (0.9)
$\mathcal{K}_{1,32}^{\mathbb{P}}$	-0.498 (0.46)	-0.576 (0.75)	-0.625 (0.22)	0.319 (0.086)	-1.17 (1.5)
$\mathcal{K}_{1,33}^{\mathbb{P}}$	-1.18 (0.87)	-0.556 (0.65)	-1.41 (0.51)	-0.131 (0.10)	-0.0376 (0.066)
$\mathcal{K}_{1,11}^{\mathbb{Q}}$	-1.15 (0.060)	-0.538 (0.016)	-0.531 (0.013)	-1.61 (0.18)	-0.573 (0.055)
$\mathcal{K}_{1,12}^{\mathbb{Q}}$	1.78 (0.077)	0	0	1.18 (0.17)	0.000 (0.067)
$\mathcal{K}_{1,13}^{\mathbb{Q}}$	1.60 (0.18)	0	0	0	0
$\mathcal{K}_{1,21}^{\mathbb{Q}}$	0.128 (0.020)	-0.569 (0.038)	-2.21 (0.46)	1.02 (0.15)	1.43 (0.17)
$\mathcal{K}_{1,22}^{\mathbb{Q}}$	-0.405 (0.010)	-0.337 (0.050)	-1.00 (0.075)	-1.42 (0.092)	-2.38 (0.093)
$\mathcal{K}_{1,23}^{\mathbb{Q}}$	-0.414 (0.023)	-0.123 (0.022)	-1.92 (0.27)	0	0
$\mathcal{K}_{1,31}^{\mathbb{Q}}$	-0.135 (0.020)	-0.760 (0.14)	-0.926 (0.12)	-0.309 (0.035)	0.579 (0.11)
$\mathcal{K}_{1,32}^{\mathbb{Q}}$	-0.136 (0.042)	-2.39 (0.14)	-0.647 (0.093)	0.571 (0.054)	-2.12 (0.28)
$\mathcal{K}_{1,33}^{\mathbb{Q}}$	-0.289 (0.051)	-1.49 (0.12)	-1.34 (0.10)	-0.133 (5.6e-3)	-0.0487 (1.7e-3)

Table 1: Parameter Estimates: $\mathcal{K}_0^{\mathbb{P}}$, $\mathcal{K}_0^{\mathbb{Q}}$, $\mathcal{K}_1^{\mathbb{P}}$, and $\mathcal{K}_1^{\mathbb{Q}}$

This table presents the parameter estimates of $\mathcal{K}_0^{\mathbb{P}}$, $\mathcal{K}_0^{\mathbb{Q}}$, $\mathcal{K}_1^{\mathbb{P}}$, and $\mathcal{K}_1^{\mathbb{Q}}$ from equation (1) with standard errors in parentheses. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 6-month Libor, and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

Parameter	Model					
	$A_0(3)$	$A_1(3)^o$	$A_1(3)$		$A_2(3)^o$	$A_2(3)$
$H_{1,11}^1$	0	1	1		1	1
$H_{1,22}^1$	0	0.385 (0.042)	3.36	(1.7)	0	0
$H_{1,33}^1$	0	2.58e-8 (0.046)	0.000	(0.049)	1.12e-3 (1.2e-3)	0.511 (0.40)
$H_{1,22}^2$	0	0	0		1	1
$H_{1,33}^2$	0	0	0		1.00e-5 (1.6e-3)	0.438 (0.36)
ρ_0	-0.193 (2.6)	0.0726 (0.035)	1.04e-3	(0.14)	0.0117 (0.080)	0.289 (0.14)
$\rho_{1,1} \times 100$	1.28 (0.026)	0.0225 (0.035)	0.0131	(0.039)	0.184 (0.033)	1.01 (0.25)
$\rho_{1,2} \times 100$	0.845 (0.072)	0.140 (0.037)	0.0772	(0.022)	-0.385 (0.043)	1.43 (0.23)
$\rho_{1,3} \times 100$	0.000 (0.15)	0.865 (0.053)	1.05	(0.10)	2.33 (0.35)	0.596 (0.12)

Table 2: Parameter Estimates: H_1 , ρ_0 , and ρ_1

This table presents the parameter estimates of H_1 , ρ_0 , and ρ_1 from equation (1) with standard errors in parentheses. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor, and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
6 Month	7.1	7.1	6.8	7.1	6.8
1 Year	9.9	9.9	9.3	10.0	9.3
3 Year	4.1	4.1	4.5	4.1	4.5
4 Year	5.3	5.2	6.3	5.2	6.2
5 Year	5.2	5.2	6.7	5.2	6.6
7 Year	3.8	3.8	5.5	3.8	5.3

Table 3: Pricing Errors in BPS for Swap Implied Zeros

This table shows the root mean squared pricing errors in basis points for yields on swap implied zeros. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

$A_2(3)$ models that we estimate without using options. Similarly, there is little difference between the $A_1(3)^o$ and $A_2(3)^o$ models that we estimate with options. The use of options to estimate the $A_1(3)^o$ and $A_2(3)^o$ models has only a small effect on the models' fits to the cross-section of zero-coupon swap rates with different maturities. Including options improves the fit (relative to models that we estimate without using options) by less than a basis point at the short end of the yield curve (up to 1 year) and worsens the fit by slightly more than a basis point at the long end of the yield curve (beyond 1 year).

Table 4 displays the root mean squared pricing errors (as a percentage of the current market price) for at-the-money caps with various maturities. For all of the models, the percentage pricing errors are worst for 1-year caps and decline as the maturity of the cap increases. The poor fit for short maturity caps, as well as the weaker fit for short maturity yields, may be due to the findings in Dai and Singleton (2002) and Piazzesi (2005) which suggest that a fourth factor is required to capture the short end of the yield curve. Additionally, as Piazzesi (2005) documents, jumps play an important role for short maturities and therefore adding jumps may improve the model along this dimension. We choose to implement more parsimonious 3-factor models because we are primarily interested in predicting changes in long-term interest rates. Amongst the models that we estimate without including options, the $A_2(3)$ model provides the best fit to the cross-section of at-the-

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
1 Year	33.6	32.4	36.4	33.3	35.5
2 Year	19.9	19.0	14.6	16.9	14.4
3 Year	18.9	18.0	10.9	15.7	10.9
4 Year	17.3	17.3	9.6	14.2	9.6
5 Year	16.3	17.0	9.2	13.3	9.0
7 Year	14.3	16.1	8.6	11.7	8.3
10 Year	13.4	15.8	9.2	11.0	8.9

Table 4: Relative Pricing Errors in % for At-the-Money Caps

This table shows the root mean squared relative pricing errors in % for at-the-money caps. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 1-, 3-, 4-, 5-, and 7-year zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

money cap prices. The $A_1(3)^o$ and $A_2(3)^o$ models have slightly larger relative pricing errors for 1-year caps than their $A_1(3)$ and $A_2(3)$ counterparts that we estimate without options. However, the relative pricing errors for caps with longer maturities are considerably lower when we include caps in estimation. For example, the root mean squared relative pricing error for at-the-money 5-year caps is 17% in the $A_1(3)$ model and 9.2% in the $A_1(3)^o$ model. Similarly, the root mean squared relative pricing error for at-the-money 5-year caps is 13.3% in the $A_2(3)$ model and 9.0% in the $A_2(3)^o$ model. The relative pricing errors for the $A_2(3)^o$ model are slightly better than those for the $A_1(3)^o$.

The pricing errors for caps from the $A_1(3)^o$ and $A_2(3)^o$ models that we estimate with options compare favorably with the pricing errors that have been reported in previous literature.¹⁴ [Driessen et al. \(2003\)](#) estimate a 3-factor Gaussian HJM model with cap data and report absolute pricing errors

¹⁴Previous papers have also used interest rate option prices other than caps to estimate dynamic term structure models. [Umantsev \(2002\)](#) finds that pricing errors for swaptions are significantly reduced when he uses swaption prices to estimate affine term structure models. [Bikbov and Chernov \(2010\)](#) use Eurodollar options to estimate term structure models with constant, stochastic, and unspanned stochastic volatility. They find that only stochastic volatility models (such as our $A_1(3)^o$ and $A_2(3)^o$ models) can reconcile both option prices and the term structure of interest rates.

(averaged across maturities) of 22.2% of cap market prices. [Li and Zhao \(2006\)](#) estimate a 3-factor quadratic term structure model and find that the root mean squared percentage pricing error for 5-year at-the-money caps is 10.4%. [Jagannathan et al. \(2003\)](#) estimate a 3-factor CIR model and find that the mean absolute pricing errors are 36.62 basis points for 5-year caps compared to a mean market price of 284.84 basis points (the results are similar for caps with other maturities). [Longstaff et al. \(2001\)](#) estimate a 4-factor string market model using swaptions and find that it overprices caps. They report that the mean percentage valuation error for 5-year caps is 5.665% and ranges from a minimum of -2.385% to a maximum of 38.071%. All of these papers report larger pricing errors for shorter-dated caps. Although not reported here, the $A_1(3)^o$ and $A_2(3)^o$ models that we estimate with caps also provide an excellent fit to the prices of at-the-money swaptions.

The pricing performance of the models estimated with and without options provides a key insight for our main objective of analyzing the predictability of excess returns. All of the models have very similar pricing errors for the cross-section of yields, which suggests that the objective function may be very flat along this dimension. Put another way, a variety of risk-neutral models can give very similar implications for bond prices and swap rates. However, when we consider cap prices there is a stark difference between models that had similar pricing errors for yields. The use of option prices in estimation provides both lower option pricing errors and more efficient estimates of the risk-neutral \mathbb{Q} parameters. In [Section 4](#), we show that this increased efficiency is also beneficial for predicting bond returns.

Unlike prices, conditional volatility is not directly observed and therefore it must be estimated.¹⁵ For estimates of conditional volatility based on historical data we use an exponential weighted moving average (EWMA) with a 26-week half-life, and also estimate an EGARCH(1,1) for each zero-coupon swap rate maturity.

[Figure 1](#) plots the conditional volatility of 10-year zero-coupon swap rates from the term structure models against our estimates of conditional volatility that use historical data. [Table 5](#) provides the correlation between the

¹⁵Implied volatilities from cap prices are forward looking and directly observable. However, in the case of models with stochastic volatility, the market prices of risk may cause the implied volatilities from cap prices to differ from the actual conditional volatility.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
6 Month	0	19.2	28.9	39.1	30.1
1 Year	0	50.8	56.3	58.3	52.9
2 Year	0	75.0	77.0	63.2	66.9
3 Year	0	83.0	81.5	39.0	70.6
4 Year	0	84.4	81.4	15.4	71.9
5 Year	0	84.1	79.3	-2.6	69.3
7 Year	0	84.3	77.4	-21.2	66.4
10 Year	0	82.0	75.0	-26.7	61.6

Table 5: Correlation between model and EGARCH volatility

This table shows the correlation between model-implied one-week volatilities and EGARCH(1,1) volatility estimates. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

conditional volatility in the pricing model and the EGARCH(1,1) estimates of conditional volatility.

The conditional volatility of all swap rates is constant in the $A_0(3)$ model and therefore these models cannot capture any time-series variation. Over our sample period, the average level of conditional volatility in the $A_0(3)$ model for zero-coupon swap rates with different maturities is slightly below our estimates based on historical data.

In the stochastic volatility models, for maturities beyond 1 year, the conditional volatility of zero-coupon swap rates in the $A_1(3)$, $A_1(3)^o$, and $A_2(3)^o$ models are all highly correlated with the EGARCH estimates of conditional volatility. The correlations are highest in the $A_1(3)$ model, followed closely by the $A_1(3)^o$ model, and then by the $A_2(3)^o$ model. The conditional volatility in the $A_2(3)$ model is positively correlated with the estimates of conditional volatility for maturities up to 4 years, but negatively correlated for maturities beyond 4 years. This result occurs because there are two stochastic volatility factors that jointly drive the level and volatility of yields in the $A_2(3)$ model. One factor is closely correlated with the volatility of short maturity yields. However, without the additional discipline in estimation that is pro-

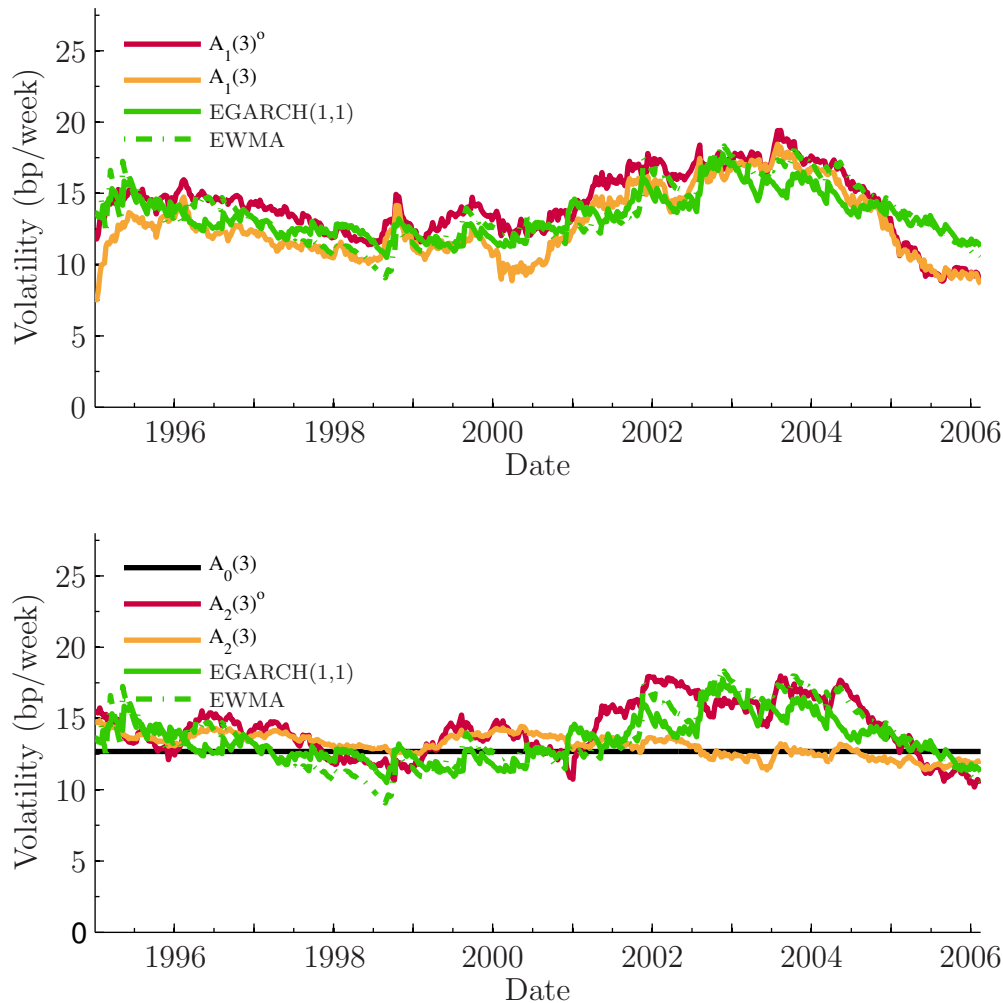


Figure 1: Realized Volatility of 10-Year Zero-Coupon Swap Rate

These figures plot weekly model conditional volatility of the 10-year zero coupon swap rate against estimates of conditional volatility based on historical data: an exponential weighted moving average (EWMA) with a 26-week half-life and an EGARCH(1,1). The top plot shows the conditional volatility in the $A_1(3)$ and $A_1(3)^\circ$ models. The bottom plot shows the conditional volatility in the $A_0(3)$, $A_2(3)$, and $A_2(3)^\circ$ models. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^\circ$ and $A_2(3)^\circ$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

vided by options, the second factor is imprecisely estimated and is negatively correlated with the volatility of long maturity yields. The $A_1(3)$ and $A_2(3)$ models match the average level of conditional volatility for the EGARCH and EWMA estimates. For the 6-month zero-coupon rate, the volatility match is worse, though still positively related. It is very similar between the $A_1(3)$ and $A_1(3)^o$ models, and between the $A_2(3)$ and $A_2(3)^o$ models.

In summary, none of the models match the conditional volatility of short-term interest rates, but the $A_1(3)$, $A_1(3)^o$, and $A_2(3)^o$ models capture the conditional volatility of long-term interest rates. The failure to match the conditional volatility of short term interest rates occurs mainly during periods of high volatility (as estimated by the EGARCH specification). Again, one could use a fourth latent factor with jumps to better capture these dynamics.

4. Predictability of Excess Returns

In this section we examine how well the risk premiums that we estimate are able to predict excess returns for long-term swaps. The excess return on a τ -maturity bond that is purchased at time t and sold at time $t + \Delta t$ is defined as

$$r_{t,\Delta t}^{e,\tau} := \underbrace{\ln(P_{t+\Delta t}^{t+\tau}/P_t^{t+\tau})/\Delta t}_{\text{bond return}} - \underbrace{\ln(1/P_t^{t+\Delta t})/\Delta t}_{\text{risk-free return}}. \quad (17)$$

In an affine term structure model this risk premium, or expected excess return, is given by

$$\mathbb{E}_t[r_{t,\Delta t}^{e,\tau}] = \frac{1}{\Delta t} \left\{ \begin{array}{l} A(\tau - \Delta t) + B(\tau - \Delta t) \cdot \mathbb{E}_t[X_{t+\Delta t}] \\ - [A(\tau) + B(\tau) \cdot X_t] + A(\Delta t) + B(\Delta t) \cdot X_t \end{array} \right\}, \quad (18)$$

where A and B satisfy the Riccati ODEs in equation (7). To measure how well our estimated risk premiums predict excess returns, we compute the following modified R^2 statistic,

$$R^2 = 1 - \frac{\text{mean} \left[(r_{t,\Delta t}^{e,\tau} - \mathbb{E}_t[r_{t,\Delta t}^{e,\tau}])^2 \right]}{\text{var} [r_{t,\Delta t}^{e,\tau}]}. \quad (19)$$

The mean (sample average) and variance in equation (19) are computed using the returns beginning in each of the 483 weeks in our sample (there is

overlap in returns with a horizon longer than a week). To emphasize, we compute our modified R^2 statistic by directly comparing realized excess returns to the model-implied expected returns computed from equation (18). Our modified statistic differs from the R^2 computed from a standard regression that uses linear projection and ignores the cross-sectional consistency of the no arbitrage model.

Table 6 presents the R^2 statistics with bootstrapped t-statistics for 3-month excess returns for the period from January 1995 to February 2006 that was used to estimate the model.¹⁶ We utilize the following semi-nonparametric procedure to bootstrap standard errors. For each bootstrap simulation, we choose a random week in the sample to initialize the state variable. We then use the estimated model parameters to simulate a 483-week sample of the state vector. Due to the lack of an economic model for the measurement errors, we complete the bootstrap procedure by using the block bootstrap with blocks of length 12 to generate pricing errors (for an overview, see Chernick, 2001). The reported bootstrapped t-statistics correspond to the reported R^2 s divided by the standard deviation of the bootstrapped R^2 s.

For comparison, in Table 6 we also provide this R^2 statistic for three versions of the regressions of excess returns on forward rates as performed in Cochrane and Piazzesi (2005). For each τ -year zero-coupon swap rate ($\tau = 2, 3, 4, 5$), Cochrane and Piazzesi (2005) regress

$$r_{t,\Delta t}^{e,n} = \beta_0^n + \beta_1^n Y_t^1 + \beta_2^n F_t^2 + \beta_3^n F_t^3 + \beta_4^n F_t^4 + \beta_5^n F_t^5 + \varepsilon_t^n, \quad (20)$$

where $F_t^n := n Y_t^{(n)} - (n-1) Y_t^{(n-1)}$ is the 1-year forward rate at time t between $t+n-1$ and $t+n$. In Table 6, we denote the regression results from equation (20) by CP₅. CP₁₀ denotes the corresponding regression results using one year forward rates up to 10 years.

Amongst the three models that we estimate without options, the $A_0(3)$ model is best at predicting 3-month excess returns for zero-coupon swaps,

¹⁶For ease of exposition, we focus on 3-month holding period returns across maturities. Roughly speaking, for shorter (longer) holding periods the R^2 s are all lower (higher) than for the 3-month holding period, with similar patterns across maturities. This result is consistent with a higher signal-to-noise ratio for shorter holding periods and is also consistent with the empirical results in Cochrane and Piazzesi (2005) and Diebold and Li (2006), among others. See also Boudoukh et al. (2008).

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	CP_5	CP_{10}
2 Yr	0.7 (9.4)	0.9 (9.6)	3.5 (9.2)	-0.2 (9.4)	3.4 (9.2)	19.8	22.8
3 Yr	6.5 (9.1)	4.0 (9.6)	8.6 (8.8)	2.7 (9.4)	9.7 (8.5)	20.1	22.2
4 Yr	10.2 (8.9)	5.9 (9.6)	11.7 (8.5)	4.0 (9.5)	13.0 (8.3)	20.1	21.6
5 Yr	12.5 (8.7)	7.0 (9.4)	13.5 (8.6)	4.6 (9.6)	14.6 (8.2)	19.7	21.2
6 Yr	13.8 (8.6)	7.5 (9.3)	14.4 (8.4)	4.9 (9.4)	15.5 (8.1)	19.0	20.3
7 Yr	14.9 (8.6)	7.9 (9.2)	15.1 (8.2)	5.2 (9.5)	16.0 (8.1)		18.5
8 Yr	15.4 (8.4)	8.0 (9.2)	15.4 (8.4)	5.2 (9.5)	16.2 (8.1)		17.9
9 Yr	15.9 (8.4)	8.0 (9.0)	15.6 (8.4)	5.3 (9.6)	16.3 (8.0)		17.5
10 Yr	16.1 (8.3)	8.0 (9.1)	15.5 (8.3)	5.3 (9.3)	16.2 (8.0)		17.1

Table 6: Predictability of Excess Returns (R^2 s)

This table presents R^2 s obtained from overlapping weekly projections of 3-month realized zero-coupon swap rate returns, for different maturities, on model implied returns. Regressions are based on overlapping data. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^o$ and $A_2(3)^o$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

	$A_1(3)^o - A_1(3)$	$A_2(3)^o - A_2(3)$
2 Yr	2.6 [-11.9, 7.4]	3.6 [-1.7, 9.1]
3 Yr	4.6 [0.1, 9.4]	7.0 [1.6, 12.8]
4 Yr	5.8 [1.4, 10.7]	9.0 [3.8, 14.8]
5 Yr	6.5 [-0.5, 11.8]	10.0 [5.0, 15.8]
6 Yr	6.9 [2.5, 11.7]	10.6 [5.4, 16.4]
7 Yr	7.2 [2.7, 12.3]	10.8 [5.8, 16.7]
8 Yr	7.4 [3.0, 12.4]	11.0 [5.9, 16.8]
9 Yr	7.6 [3.2, 12.3]	11.0 [6.2, 16.6]
10 Yr	7.5 [3.2, 12.2]	10.9 [6.3, 16.4]

Table 7: Difference in R^2 's

This table presents the difference in R^2 's for 3-month realized zero-coupon swap rate returns using the model-implied expected returns. The table presents the differences between the $A_1(3)^o$ and $A_1(3)$ models, and between the $A_2(3)^o$ and $A_2(3)$ models. Bootstrapped 95% confidence intervals are presented below in parentheses and are computed using the bootstrap procedure described above. We estimated the $A_1(3)$ and $A_2(3)$ models by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month LIBOR and 1-, 3-, 4-, 5-, and 7-year zeros with error. We estimated the $A_1(3)^o$ and $A_2(3)^o$ models with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

followed by the $A_1(3)$ model and then by the $A_2(3)$ model. When we include options in estimation the results are reversed. The $A_2(3)^o$ model provides the best prediction results, followed by the $A_1(3)^o$ model and then the $A_0(3)$ model (the $A_0(3)$ model has slightly higher R^2 s for the 9- and 10-year maturities). Moreover, when we include options in estimation, the R^2 s are much closer in magnitude to those obtained from the regressions in [Cochrane and Piazzesi \(2005\)](#). The regressions in [Cochrane and Piazzesi \(2005\)](#) are designed to match only excess returns and so they serve as an upper bound for the level of predictability. Notice that the worst predictive performance is for short maturities, which is the same as the pricing results in [Tables 3 and 4](#). In the $A_2(3)$ model, the R^2 is even negative for the 3-month excess return on the 2-year zero-coupon swap, which indicates that the model's predictions are worse than the sample average of returns. We attribute this poorer performance for short-maturity zero-coupon swaps to the same reasons we discussed in [Section 3](#) on the fit to prices.

[Figure 2](#) plots the 3-month expected excess return on a 5-year zero-coupon bond for each of the models that we estimate. Consistent with the results in [Tables 6 and 7](#), the expected excess returns for the $A_1(3)^o$ and $A_2(3)^o$ models that we estimate using options are very similar to those for the constant volatility $A_0(3)$ model (which was the preferred model in [Duffee, 2002](#) and [Cheridito et al., 2007](#)). The expected excess returns for the $A_1(3)$ and $A_2(3)$ models are very similar to each other, but different from their counterparts that we estimate using options.

To summarize, the risk premiums that we estimate using interest rate cap prices are better able to predict excess returns for long term swaps. The question remains, what elements of the risk premium do options help to identify? To address this question, [Table 8](#) highlights the properties of the risk-neutral distribution which differ when we use options to estimate the models.

The risk-neutral long-run means ($\theta_{3m}^{\mathbb{Q}}$, $\theta_{2y}^{\mathbb{Q}}$, and $\theta_{10y}^{\mathbb{Q}}$) of the 3-month, 2-year, and 10-year zero-coupon yields range between 10.4% and 13% for the models that are estimated without options, but hover around 8% for the $A_1(3)^o$ and $A_2(3)^o$ models that are estimated with options. For comparison, the maximum values of the 3-month, 2-year, and 10-year zero-coupon swap rates over our sample period are 6.8%, 8%, and 8.2%, respectively, and the mean values are 4.2%, 4.8%, and 5.9%. In each model, the third eigenvector

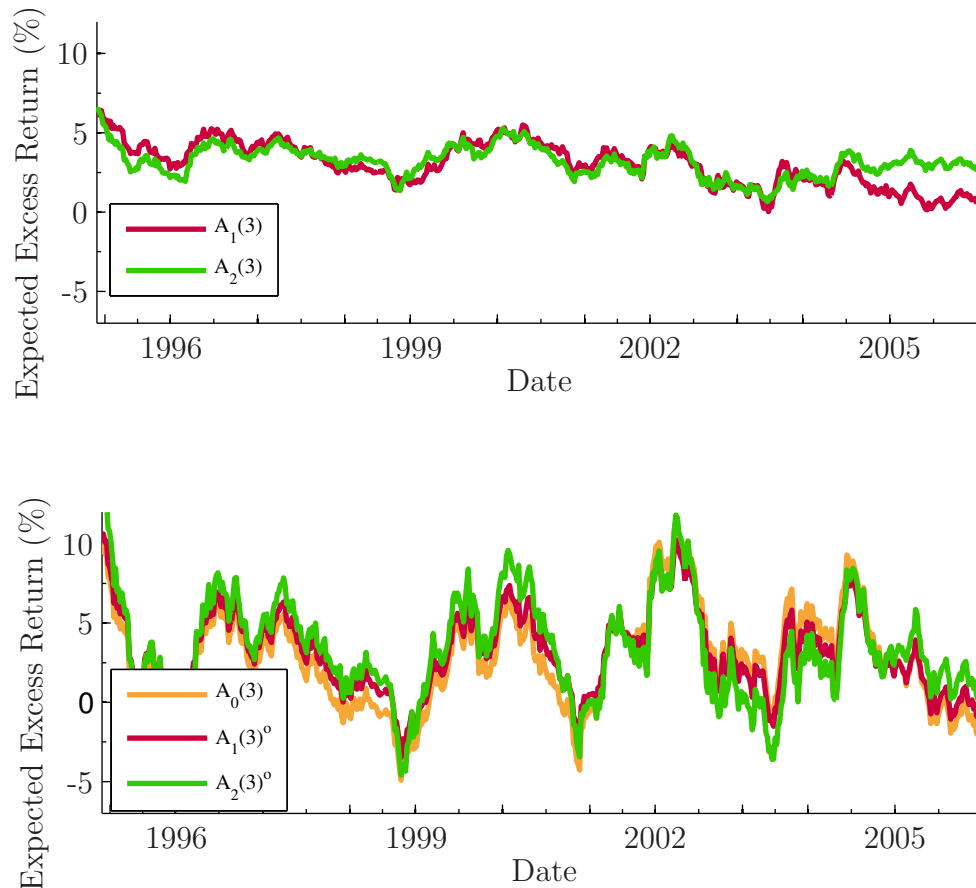


Figure 2: Quarterly Expected Excess Return – 5-Year Zero Coupon

This figure plots the quarterly expected excess return on a 5-year zero-coupon bond (annualized) for each of the models that we estimate. The top plot shows the expected excess return for the $A_1(3)$ and $A_2(3)$ models that we estimated without using options. The bottom plot shows the expected excess return for the $A_0(3)$ model that we estimated without using options, and the $A_1(3)^o$ and $A_2(3)^o$ models that we estimated using options.

Parameter	Estimate				
	$\mathbb{A}_0(3)$	$\mathbb{A}_1(3)$	$\mathbb{A}_1(3)^\circ$	$\mathbb{A}_2(3)$	$\mathbb{A}_2(3)^\circ$
$\theta_{3m}^{\mathbb{Q}}$	11.10	10.50	8.02	13.00	8.14
$\theta_{2y}^{\mathbb{Q}}$	11.10	10.50	8.02	13.00	8.14
$\theta_{10y}^{\mathbb{Q}}$	11.00	10.40	7.86	12.70	7.98
half-life $_1^{\mathbb{Q}}$	0.53	0.49	0.44	0.49	0.48
half-life $_2^{\mathbb{Q}}$	1.02	1.12	1.22	1.06	1.16
half-life $_3^{\mathbb{Q}}$	23.30	22.40	8.09	31.40	8.46
eigenvector $_{3,3m}$	0.61	0.61	0.66	0.61	0.65
eigenvector $_{3,2y}$	0.59	0.59	0.61	0.59	0.61
eigenvector $_{3,10y}$	0.53	0.53	0.45	0.53	0.45

Table 8: Mean Reversion and Long-Run Means

This table reports the risk-neutral long-run means ($\theta_{3m}^{\mathbb{Q}}$, $\theta_{2y}^{\mathbb{Q}}$, and $\theta_{10y}^{\mathbb{Q}}$) of the 3-month, 2-year, and 10-year zero-coupon yields. It also reports the half-life of the eigenvectors of the mean reversion matrix under \mathbb{Q} , as well as the elements of the third eigenvector. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 1-, 3-, 4-, 5-, and 7-year zeros with error. The $A_1(3)^\circ$ and $A_2(3)^\circ$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were priced with error.

of $\mathcal{K}^{\mathbb{Q}}$ acts as a level factor that is almost equally weighted across the three yields. For the models that are estimated without options, the half-life of shocks to this level factor ranges from 22.4 years to 31.4 years, while it is just 8.09 years in the $A_1(3)^{\circ}$ model and 8.46 years in the $A_2(3)^{\circ}$ model. That is, for the models that are estimated without options, the risk-neutral expectation of future yields is high, but shocks to the level of yields are very persistent. By contrast, for the $A_1(3)^{\circ}$ and $A_2(3)^{\circ}$ models, under the risk-neutral measure the level of yields reverts more quickly to a lower long-run mean (i.e. shocks are less persistent and die off more quickly). Both of these combinations produce an upward sloping yield curve consistent with the data. However, the faster mean reversion to a lower long-run mean fits the joint swaps and options data best. Although heuristic, we also feel that the lower long-run means represent a more economically plausible level of risk aversion with a rate of mean reversion that is on the order of business cycle frequency. For example, [Gurkaynak et al. \(2005\)](#) find difficulty reconciling the business cycle frequency variation (under \mathbb{P}) in macroeconomic variables with the long frequency variation implicit in long maturity bond yields (under \mathbb{Q}). To some extent, our findings that the options-based estimates have higher rates of mean reversion (under \mathbb{Q}) suggest that information in option prices partially attenuates this difference.

5. Factor Analysis

[Litterman and Scheinkman \(1991\)](#) find that most of the variation in bond returns can be explained by the level, slope, and curvature of the yield curve. In this section we examine how the expected excess returns in each model depend on the level, slope, and curvature of interest rates (which are, in turn, affine functions of the underlying states in the model). In each model that we estimate we can write

$$\begin{aligned} \mathbb{E}_t [r_{t,\Delta t}^{e,\tau}] &= \text{CONSTANT} + \text{LEVEL} \times Y_t^{(0.25)} + \text{SLOPE} \times \left(Y_t^{(10)} - Y_t^{(0.25)} \right) \\ &\quad + \text{CURVATURE} \times \left(2 \cdot Y_t^{(2)} - Y_t^{(10)} - Y_t^{(0.25)} \right), \end{aligned} \quad (21)$$

where $Y_t^{(0.25)}$, $Y_t^{(2)}$, and $Y_t^{(10)}$ are the 3-month, 2-, and 10-year zero-coupon swap rates, respectively. For ease of exposition, we focus on the 10-year return decomposition for a holding period (Δt) of 3 months. Results are

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
CONSTANT	-0.224	-0.129	-0.159	0.012	-0.127
LEVEL	3.157	2.920	2.672	0.683	2.455
SLOPE	9.030	3.263	6.631	1.173	6.443
CURVATURE	3.347	0.680	3.524	2.080	6.231

Table 9: Composition of Expected Excess Return

This table provides the values for **CONSTANT**, **LEVEL**, **SLOPE**, and **CURVATURE** for each of the models that we estimate, where the 3-month expected excess return on a 10-year zero-coupon bond is

$$\begin{aligned} \mathbb{E}_t \left[r_{t,1}^{e,5} \right] &= \text{CONSTANT} + \text{LEVEL} \times Y_t^{(0.25)} + \text{SLOPE} \times \left(Y_t^{(10)} - Y_t^{(0.25)} \right) \\ &\quad + \text{CURVATURE} \times \left(2 \cdot Y_t^{(2)} - Y_t^{(10)} - Y_t^{(0.25)} \right). \end{aligned}$$

similar for other moderate and long maturities. Table 9 provides the values for **CONSTANT**, **LEVEL**, **SLOPE**, and **CURVATURE** for each of the models that we estimate. Comparing the $A_1(3)$ and $A_1(3)^o$ models, we see that in general the options model gives more weight to **SLOPE** and **CURVATURE** (3.26 versus 6.63 and 0.67 vs 3.52, respectively.) The weight on level is similar across these models. Regarding the $A_2(3)$ and $A_2(3)^o$ models, we see a similar pattern across **SLOPE** and **CURVATURE**: including options results in a larger weight on these two variables (0.68 versus 2.45 and 2.08 versus 6.23, respectively.) Note that although the differences are of similar magnitude, **SLOPE** has roughly three times the standard deviation of **CURVATURE**, therefore we can conclude that including options in model estimation primarily helps to identify the risk premium (or expected excess return) that is associated with the slope of the yield curve and, to a lesser degree, the curvature of the yield curve.

The $A_0(3)$ model shows a roughly similar pattern to the models that we estimate with options, with comparable values of **LEVEL** and higher values of **SLOPE** and **CURVATURE**. This result is consistent with our findings that, as a group, the $A_0(3)$, $A_1(3)^o$ and $A_2(3)^o$ outperformed the $A_1(3)$ and $A_2(3)$ models with respect to forecasting bond returns.

[Dai and Singleton \(2002\)](#) present two additional challenges for dynamic term structure models that are related to risk premia in fixed income markets. The first challenge, which [Dai and Singleton \(2002\)](#) refer to as the

linear projection of yields, or LPY(I), is to match the pattern of violations of the expectations hypothesis that is documented in [Fama and Bliss \(1987\)](#) and [Campbell and Shiller \(1991\)](#). These papers regress excess returns on the slope of the yield curve and find that the regression coefficient is not 1, but instead is negative (and more so for longer maturities). Dai and Singleton’s second related challenge, LPY(II), states that when excess returns are adjusted by model-implied risk premia, the expectations hypothesis should be restored and the regression coefficient on the slope of the yield curve should be 1.

Specifically, [Fama and Bliss \(1987\)](#) and [Campbell and Shiller \(1991\)](#) perform the following regression

$$Y_{t+\Delta t}^{n-\Delta t} - Y_t^n = \Phi_{0,n} + \Phi_{1,n} \frac{\Delta t}{n - \Delta t} (Y_t^n - Y_t^{\Delta t}) + \varepsilon_t^n, \quad (22)$$

and find that the regression coefficients, $\hat{\Phi}_{1,n}$, are increasingly negative for larger maturities n . [Figure 3](#) provides the LPY(I) regression results for the models that we estimate.¹⁷ [Dai and Singleton \(2002\)](#) find that only the $A_0(3)$ model matches LPY(I). We also find that, when not using options data, only the $A_0(3)$ model with constant volatility matches LPY(I) while the models with stochastic volatility fail to replicate this feature of the data. In contrast, we find that the models with stochastic volatility, when estimated jointly with options data, are statistically consistent with the empirically observed slope coefficients. For all of the models, the observed slope coefficients lie within a simulated 95% confidence interval. However, the point estimates in the $A_0(3)$, $A_1(3)^o$, and $A_2(3)^o$ models are much closer to the observed regression coefficients in the data. In unreported results, we also found that, when adjusted for small-sample bias, the $A_1(3)^o$ and $A_2(3)^o$ models that we estimate with options also satisfy LPY(II). Our results differ from the findings in both [Duffee \(2002\)](#) and [Dai and Singleton \(2002\)](#) that dynamic term structure models with constant volatility were superior to those with stochastic volatility, in terms of consistency in predicting yields. Instead, we

¹⁷The mean regression coefficients for each model were generated using 1,000 simulations of an Euler approximation of the stochastic volatility models. In the case of the $A_0(3)$, the confidence interval can in fact be computed exactly in closed form. The stochastic volatility models do not yield closed form expressions for the confidence intervals since closed form transition densities are not known for these processes.

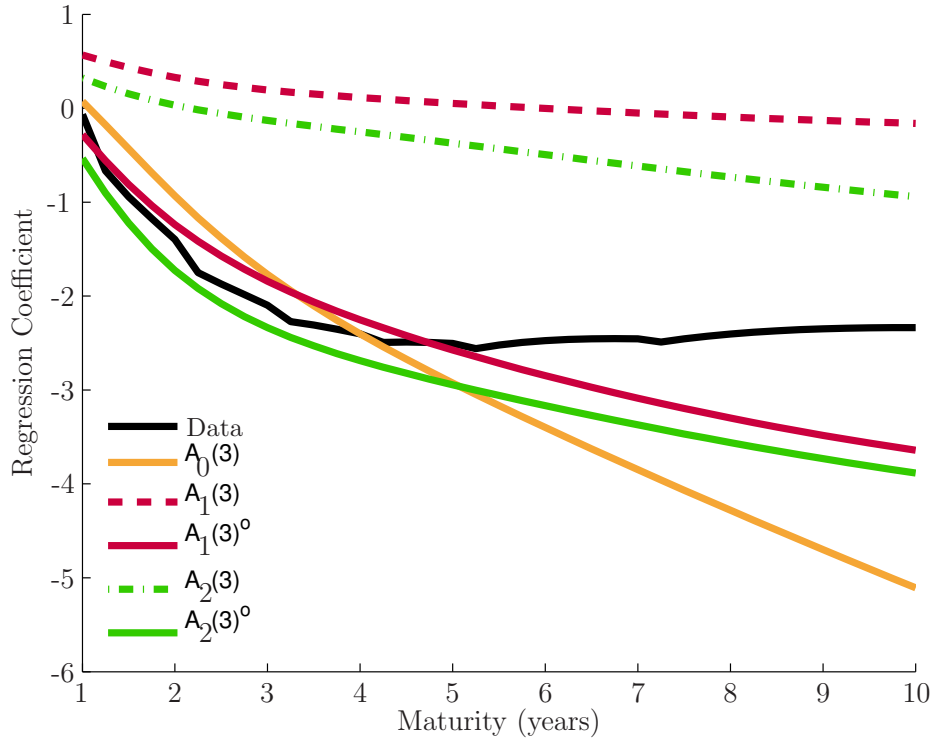


Figure 3: Regression Coefficients from Linear Projection on Yields

This figure shows the regression coefficients of the Campbell-Shiller regression $Y_{t+0.25}^{(n-0.25)} - Y_t^n$ on slope, $(Y_t^n - Y_t^{(0.25)})/(n - 0.25)$. The model values are simulated mean regression coefficients. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and pricing 6-month Libor and 1-, 3-, 4-, 5-, and 7-year swap zeros with error. The $A_1(3)^\circ$ and $A_2(3)^\circ$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money cap were priced with error.

find that using options data for estimation allows the models with stochastic volatility to maintain similar or superior performance along this dimension.

6. Conclusion

Interest rate options may contain information about the risk premium in long-term interest rates because their prices are sensitive to the volatility and market prices of the underlying interest rates. We use the time series of interest rate cap prices and swap rates to estimate 3-factor affine term structure models. The risk premiums that we estimate using interest rate option prices are better able to predict excess returns for long-term swaps over short-term swaps. We show that including options reduces the estimate of the risk-neutral long-run mean of yields and increases the estimated rate of mean reversion to a more economically plausible level. We also show that including options helps us to identify the portion of the risk premium that is associated with the slope and curvature of the yield curve. With a correction for a small sample bias, the models with stochastic volatility that we estimate with options also capture the failure of the expectations hypothesis and match regressions of returns on the slope of the yield curve.

Acknowledgements

We thank Chris Armstrong, Snehal Banerjee, David Bolder, Mikhail Chernov, Anna Cieslak, Darrell Duffie, Peter Feldhütter, René Garcia, Chris Jones, Yaniv Konchitchki, Peter Reiss, David Runkle, and seminar participants at the American Finance Association meetings in Chicago, the Northern Finance Association meetings in Vancouver, the SAFE 2010 Term Structure Conference in Verona, the University of Waterloo, the Bank of Canada Fixed Income conference, Barclays Global Investors, and Stanford GSB. We are very grateful to Ken Singleton for many discussions and comments. Almeida gratefully acknowledges financial support from CNPq-Brazil and Graveline acknowledges support from a Carlson School Dean's Small Research Grant.

- Ahn, D.H., Dittmar, R.F., Gallant, A.R., 2002. Quadratic term structure models: Theory and evidence. *The Review of Financial Studies* 15, 243–288.
- Aït-Sahalia, Y., Lo, A.W., 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94, 9–51.
- Aït-Sahalia, Y., Wang, Y., Yared, F., 2001. Do option markets correctly price the probabilities of movement of the underlying asset? *Journal of Econometrics* 102, 67–110.
- Andersen, T.G., Benzoni, L., 2008. Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models. SSRN eLibrary .
- Bakshi, G.S., Carr, P.P., Wu, L., 2008. Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. *Journal of Financial Economics* 87, 132–156.
- Bikbov, R., Chernov, M., 2010. Yield curve and volatility: Lessons from eurodollar futures and options. *Journal of Financial Econometrics* forthcoming.
- Boudoukh, J., Richardson, M., Whitelaw, R.F., 2008. The myth of long-horizon predictability. *Review of Financial Studies* 21, 1577–1605.
- Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: A bird’s eye view. *Review of Economic Studies* 58, 495–514.
- Chen, R.R., Scott, L., 1993. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income* 3, 14–31.
- Cheng, P., Scaillet, O., 2007. Linear-quadratic jump-diffusion modeling. *Mathematical Finance* 17, 575–598.
- Cheridito, P., Filipovic, D., Kimmel, R.L., 2007. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics* 83, 123–170.
- Chernick, M., 2001. *Bootstrap Methods: A Guide for Practitioners and Researchers*. Wiley-Interscience. 3rd edition.

- Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics* 56, 407–458.
- Cochrane, J.H., Piazzesi, M., 2005. Bond risk premia. *American Economic Review* 95, 138–160.
- Collin-Dufresne, P., Goldstein, R.S., 2002a. Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *Journal of Finance* 57, 1685–1730.
- Collin-Dufresne, P., Goldstein, R.S., 2002b. Pricing swaptions within an affine framework. *Journal of Derivatives* 10, 1–18.
- Collin-Dufresne, P., Goldstein, R.S., Jones, C.S., 2009. Can interest rate volatility be extracted from the cross section of bond yields? *Journal of Financial Economics* 94, 47–66.
- Cox, J.C., Jonathan E. Ingersoll, J., Ross, S.A., 1985. A theory of the term structure of interest rates. *Econometrica* 53, 385–408.
- Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.
- Dai, Q., Singleton, K.J., 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics* 63, 415–441.
- Dai, Q., Singleton, K.J., 2003. Term structure dynamics in theory and reality. *Review of Financial Studies* 16, 631–678.
- Davidson, R., MacKinnon, J., 1993. *Estimation and Inference in Econometrics*. Oxford University Press.
- Diebold, F., Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337–364.
- Driessen, J., Klaassen, P., Melenberg, B., 2003. The performance of multi-factor term structure models for pricing and hedging caps and swaptions. *Journal of Financial and Quantitative Analysis* 38, 635–672.

- Duarte, J., 2004. Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies* 17, 379–404.
- Duffee, G.R., 2002. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57, 405–443.
- Duffie, J.D., 2001. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey. 3rd edition.
- Duffie, J.D., Kan, R., 1996. A yield-factor model of interest rates. *Mathematical Finance* 6, 379–406.
- Duffie, J.D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.
- Eraker, B., 2004. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance* 59, 1367–1403.
- Fama, E.F., 1984a. The information in the term structure. *Journal of Financial Economics* 13, 509–528.
- Fama, E.F., 1984b. Term premiums in bond returns. *Journal of Financial Economics* 13, 529–546.
- Fama, E.F., Bliss, R.R., 1987. The information in long-maturity forward rates. *American Economic Review* 77, 680–692.
- Fisher, M., Gilles, C., 1996. Estimating exponential-affine models of the term structure. Working Paper.
- Graveline, J.J., 2008. Exchange rate volatility and the forward premium anomaly. Working Paper.
- Gurkaynak, R.S., Sack, B., Swanson, E., 2005. Stochastic volatilities and correlations of bond yields. *American Economic Review* 95, 425–436.
- Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press.
- Han, B., 2007. Stochastic volatilities and correlations of bond yields. *Journal of Finance* 62, 1491–1524.

- Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6, 327–343.
- Jackwerth, J.C., 2000. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13, 433–451.
- Jagannathan, R., Kaplin, A., Sun, S., 2003. An evaluation of multi-factor cir models using libor, swap rates, and cap and swaption prices. *Journal of Econometrics* 116, 113–146.
- Jones, C.S., 2003. The dynamics of stochastic volatility: Evidence from underlying and options markets. *Journal of Econometrics* 116, 181–224.
- Joslin, S., 2007. Pricing and hedging volatility risk in fixed income markets. Working Paper.
- Kim, D.H., 2007. Spanned stochastic volatility: A reexamination of the relative pricing between bonds and bond options. BIS Working Papers.
- Leippold, M., Wu, L., 2002. Asset pricing under the quadratic class. *The Journal of Financial and Quantitative Analysis* 37, 271–295.
- Li, H., Zhao, F., 2006. Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. *Journal of Finance* 61, 341–378.
- Liptser, R., Shiryaev, A., Aries, B., 2000. *Statistics of Random Processes I*. Springer.
- Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1, 54–61.
- Longstaff, F.A., Santa-Clara, P., Schwartz, E.S., 2001. The relative valuation of caps and swaptions: Theory and empirical evidence. *Journal of Finance* 56, 2067–2109.
- Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.
- Piazzesi, M., 2005. Bond yields and the federal reserve. *Journal of Political Economy* 113, 311–344.

- Thompson, S.B., 2008. Identifying term structure volatility from the libor-swap curve. *The Review of Financial Studies* 21, 819–854.
- Umantsev, L., 2002. *Econometric Analysis of European Libor-Based Options within Affine Term-Structure Models*. Ph.D. thesis. Stanford University.