Do Interest Rate Options Contain Information About Excess Returns?*

Caio Almeida[†] Jeremy J. Graveline[‡] Scott Joslin[§]

October 30, 2008

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.

*We thank Chris Armstrong, Snehal Banerjee, David Bolder, Mikhail Chernov, Darrell Duffie, Peter Feldhütter, René Garcia, Chris Jones, Yaniv Konchitchki, Peter Reiss, and David Runkle. We are very grateful to Ken Singleton for many discussions and comments.
[†]Graduate School of Economics, Fundação Getulio Vargas, calmeida@fgv.br
[‡]Carlson School of Management, University of Minnesota, jeremy@umn.edu
[§]MIT Sloan School of Management, sjoslin@mit.edu

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 the risk premium in long-term interest rates. 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. As a measure of the predictability of excess returns we compute the R^2 from the projection of 1-year realized zero coupon swap rate returns on model-implied expected returns. On average across different maturities, this measure of predictability almost doubles when we use options to estimate a term structure model with one stochastic volatility factor. The improvement is almost three fold 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 risk factors are quite different when the models are estimated with and without options. In the models estimated without options, long-run interest rates under the risk-neutral measure are high (10.4-13%) and have a very 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

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

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

is about 8 years). Both of these combinations have similar implications for swap rates, which makes them difficult to distinguish using only yields. 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. We also show that the magnitude of the estimated price of volatility risk is reduced when we include options.

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

[3] Swaps 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.

[4] The 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). See Fama (1984a), Fama (1984b), Fama and Bliss (1987) and Campbell and Shiller (1991).

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 economic role because they cannot be replicated using the underlying bonds or swaps. Our objective in this paper is to focus exclusively on the econometric benefits of using options to estimate the risk premium in long-term interest rates. Therefore our empirical analysis only considers term structure models in which options are redundant securities.

Other papers have used options to estimate term structure models, but they do not examine the impact of options 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 (2005) 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]

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

[5]See Collin-Dufresne and Goldstein (2002b), Collin-Dufresne et al. (2008), Andersen and Benzoni (2008), Li and Zhao (2006), Thompson (2008), Bikbov and Chernov (2005), Joslin (2007), and Kim (2007).

[6]See 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 objection distribution of returns inferred from time series data.

and historical volatilities are provided in appendices.

2 Model and Estimation Strategy

To fix notation, let P_t^T be the price at time t of a zero-coupon bond that pays \$1 at time T. If the bond is sold before it matures, say at time $t+\Delta t$, then it has an uncertain return $\ln\left(P_{t+\Delta t}^T / P_t^T\right)/\Delta t$. A bond that matures at time $t+\Delta t$ provides a certain return $-\ln\left(P_t^{t+\Delta t}\right)/\Delta t$, so the excess return on the longer dated bond is given by

$$r_{t,\Delta t}^{e,T} := \ln\left(P_{t+\Delta t}^T / P_t^T\right)/\Delta t + \ln\left(P_t^{t+\Delta t}\right)/\Delta t\,. \tag{1}$$

Our objective in this paper is to examine how well different arbitrage-free affine term structure models predict this excess return for long-term swaps.

To be specific, we examine 3-factor term structure models in which the pricing kernel follows a diffusion process of the form

$$dM_t = -M_t\, r_t\, dt - M_t\, \Lambda_t^\top d\, W_t\,, \tag{2a}$$

where

$$r_t := \rho_0 + \rho_1 \cdot X_t\,,^{7} \tag{2b}$$

$$\Lambda_t := \left(\sqrt{\Delta\left[\alpha + \beta X_t\right]}\right)^{-1}\left[\left(\mathcal{K}_0^{\mathbb{P}} - \mathcal{K}_0\right) + \left(\mathcal{K}_1^{\mathbb{P}} - \mathcal{K}_1\right) X_t\right]\,, \tag{2c}$$

and

$$dX_t = \left[\mathcal{K}_0^{\mathbb{P}} + \mathcal{K}_1^{\mathbb{P}} X_t\right] dt + \sqrt{\Delta\left[\alpha + \beta X_t\right]}\, d\, W_t\,.^{8} \tag{2d}$$

In equation (2), X_t is a 3-dimensional vector of latent factors, W_t is a 3-dimensional Brownian motion, and $\Delta\left[\cdot\right]$ denotes a square matrix with its vector argument along the diagonal. We refer to the drift of the pricing kernel, r, as the short interest rate and the volatility, Λ, as the market price of risk.[9]

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

[8]Dai and Singleton (2000) and Cheridito et al. (2007) provide restrictions required to ensure that the stochastic processes are admissable, the parameters are identified, and the physical and risk neutral measures are equivalent.

[9]We use an extended affine market price of risk introduced by Cheridito et al. (2007) as a generalization of the essentially affine market price of risk used in Duffee (2002).

In this setting, the dynamics of zero-coupon bond prices are given by

$$dP_t^T = P_t^T \left[r_t + \sigma_t^T \Lambda_t\right] dt + P_t^T \sigma_t^T dW_t \,, \tag{3}$$

where $\sigma_t^T = B(T-t)^\top \sqrt{\Delta[\alpha + \beta X_t]}$ and $B(\cdot)$ is the solution to a Riccati ODE.[10] From equation (3), the volatility of zero-coupon bond prices or yields is determined by the volatility of the risk factors. Also, the instantaneous expected excess return, or risk premium, is given by

$$\sigma_t^T \Lambda_t = B(T-t)^\top \sqrt{\Delta[\alpha + \beta X_t]}\, \Lambda_t \,, \tag{6}$$

which depends on both the volatility and market prices of the risk factors.

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 that is paid on the floating side of a swap.[11] 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 the market prices of risk, which are embedded in the pricing kernel.

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

$$C_t^N\left(\overline{C}\right) = \sum_{n=2}^{N} \mathbb{E}_t\left[\frac{M_{t+0.25n}}{M_t}\, \underbrace{0.25\left(L_{t+0.25(n-1)} - \overline{C}\right)^+}_{\text{caplet payoff}}\right], ^{12} \tag{7}$$

[10] Duffie and Kan (1996) show that zero-coupon bond prices are an exponential affine function of the factors,

$$P_t^T = e^{A(T-t)+B(T-t)\cdot X_t}\,, \tag{4}$$

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

$$\frac{d}{d\tau}B(\tau) = -\rho_1 + \mathcal{K}_1^\top B(\tau) + \frac{1}{2}\beta^\top \Delta[B(\tau)]\, B(\tau)\,, \quad B(0) = 0\,, \tag{5a}$$

$$\frac{d}{d\tau}A(\tau) = -\rho_0 + \mathcal{K}_0^\top B(\tau) + \frac{1}{2}\alpha^\top \Delta[B(\tau)]\, B(\tau)\,, \quad A(0) = 0\,. \tag{5b}$$

[11] Other papers, such 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.

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

$$1 + 0.25L_{t+0.25(n-1)} = 1/P_{t+0.25(n-1)}^{t+0.25n}. \quad (8)$$

Duffie et al. (2000) show that cap prices can be computed as a sum of inverted Fourier transforms in affine term structure models. However, when the solutions A and B to the Riccati ODEs in equation (5) are not known in closed form, direct Fourier inversion can be too computationally expensive for use in estimation. Instead, we use a more computationally efficient adaptive quadrature method that is based on Joslin (2007).

Our data, obtained from Datastream, consists of Libor rates, swap rates, and at-the-money cap implied volatilities from January 1995 to February 2006. We use 3-month Libor and the entire term structure of swap rates to bootstrap zero-coupon swap rates at 1-, 2-, 3-, 5-, 7-, and 10-years.[13] We also use at-the-money caps with maturities of 1-, 2-, 3-, 4-, 5-, 7-, and 10-years.

Our empirical analysis centers on three different model specifications with 0, 1, or 2 of the factors driving the stochastic volatility of interest rates. Following Dai and Singleton (2000), we use the notation $A_M(3)$ to denote a 3-factor affine term structure model in which M of the factors drive the stochastic volatility of interest rates.[14] We use quasi-maximum likelihood to estimate model parameters for $A_0(3)$, $A_1(3)$, and $A_2(3)$ models. We follow Chen and Scott (1993) and estimate all of the models under the assumption that the model correctly prices 3-month Libor and the 2- and 10-year zero coupon swap rates exactly and we assume that the remaining zero-coupon swap rates are priced with error.[15] 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. The parameter estimates and a detailed description of the estimation procedure are provided in Appendix A.

[13] Our bootstrap procedure assumes that forward swap zero rates are constant between observations.

[14] Dai and Singleton (2000) provide a canonical representation of $A_M(3)$ models.

[15] By assuming that a subset of securities are 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.

3 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. Recall that the risk premium, or expected excess return on a τ-maturity bond that is purchased at time t and sold at time $t + \Delta t$ is

$$\mathbb{E}_t\left[r_{t,\Delta t}^{e,\tau}\right] := \mathbb{E}_t\left[\ln\left(P_{t+\Delta t}^{t+\tau}/P_t^{t+\tau}\right)\right]/\Delta t - r_t^{\Delta t}\,. \tag{9}$$

In an affine term structure model, this risk premium is given by

$$\mathbb{E}_t\left[r_{t,\Delta t}^{e,\tau}\right] = \frac{1}{\Delta t}\left\{\begin{array}{l} A(\tau-\Delta t) + B(\tau-\Delta t)\cdot\mathbb{E}_t\left[X_{t+\Delta t}\right] \\ -\left[A(\tau)+B(\tau)\cdot X_t\right] + A(\Delta t) + B(\Delta t)\cdot X_t \end{array}\right\}. \tag{10}$$

To measure how well our estimated risk premiums predict excess returns, we compute the following statistic

$$R^2 = 1 - \frac{\text{mean}\left[\left(r_{t,\Delta t}^{e,\tau} - \mathbb{E}_t\left[r_{t,\Delta t}^{e,\tau}\right]\right)^2\right]}{\text{var}\left[r_{t,\Delta t}^{e,\tau}\right]}\,. \tag{11}$$

We then compare the R^2s for each model with the R^2s from three versions of the regressions of excess returns on forward rates as performed in Cochrane and Piazzesi (2005).

Table 1 presents the R^2 statistics and bootstrapped confidence intervals for 1 year excess returns for the period from January 1995 to February 2006 that was used to estimate the model. Amongst the three models that we estimate without options, the $A_1(3)$ model is best at predicting 1 year excess returns for zero-coupon swaps with maturities up to 5 years. For these maturities, the $A_2(3)$ model outperforms the $A_0(3)$ model. For maturities beyond 5 years, the $A_0(3)$ model is best at predicting 1 year excess returns, followed by the $A_1(3)$ model and then the $A_2(3)$ model. Table 13 in Appendix E provides the predictability results for 3 month excess returns with a shorter horizon and more independent observations. The results are qualitatively the same as those for 1 year excess returns.

When we include options in estimation, the $A_1(3)^o$ and $A_2(3)^o$ models are better able to predict 1 year excess returns relative to their $A_1(3)$ and $A_2(3)$ counterparts and the $A_0(3)$ model that we estimated without including options.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	CP_5	CP_{10}	$CP_{5,10}$
2 Yr	-71.0	-62.6	-68.6	-58.1	-66.4	36.4	45.3	36.1
	[-104.2, -42.9]	[-92.4, -38.0]	[-98.9, -42.3]	[-86.4, -33.7]	[-98.1, -40.2]	[30.8, 43.0]	[38.2, 54.4]	[28.5, 42.4]
3 Yr	-23.4	-16.2	-17.1	-14.9	-11.9	42.5	50.4	40.8
	[-46.8, -3.4]	[-37.8, 1.6]	[-38.2, 1.3]	[-36.3, 3.5]	[-32.4, 5.7]	[36.0, 48.9]	[44.0,58.3]	[32.5, 47.0]
4 Yr	-4.9	-0.7	1.8	-3.4	7.2	47.8	54.6	44.6
	[-24.7, 12.1]	[-19.3, 14.9]	[-16.0, 17.1]	[-23.0, 13.6]	[-9.7, 21.7]	[41.1, 54.0]	[47.9, 61.8]	[36.0, 51.0]
5 Yr	6.2	7.4	12.9	2.1	17.7	51.5	57.5	46.8
	[-11.4, 21.5]	[-10.6, 22.5]	[-3.4, 26.4]	[-17.3, 18.7]	[2.4, 31.0]	[45.1, 57.5]	[51.7, 64.1]	[37.4, 53.2]
6 Yr	13.3	11.5	19.7	4.8	23.7		59.0	47.0
	[-3.1, 27.2]	[-6.4, 26.7]	[4.7, 32.6]	[-14.5, 21.2]	[9.1, 36.4]		[53.1, 65.7]	[37.9, 53.6]
7 Yr	20.4	14.9	26.3	8.0	29.4		60.0	46.4
	[4.8, 33.5]	[-3.2, 30.2]	[11.8, 38.5]	[-10.8, 24.2]	[14.7, 41.7]		[53.9, 66.8]	[37.3, 52.3]
8 Yr	23.2	15.5	28.5	8.3	30.9		60.0	45.0
	[8.5, 35.7]	[-3.1, 30.8]	[14.4, 40.6]	[-11.3, 24.8]	[16.0, 43.2]		[54.4, 66.0]	[35.4, 50.6]
9 Yr	27.0	16.8	31.9	9.8	33.5		59.9	43.5
	[12.3, 39.4]	[-1.9, 32.0]	[17.8, 44.0]	[-10.3, 26.7]	[19.0, 45.5]		[53.9, 66.3]	[33.9, 49.3]
10 Yr	29.6	17.3	33.9	10.8	35.0		59.8	42.0
	[15.4, 41.6]	[-0.9, 33.0]	[20.5, 45.5]	[-9.2, 27.3]	[20.8, 47.0]		[53.7, 66.2]	[32.3, 47.3]

Table 1: Predictability of Excess Returns (R^2s in %)

This table presents R^2s obtained from overlapping weekly projections of one year realized zero coupon swap rate returns, for different maturities, on model implied returns. Bootstrapped standard errors are presented below in parentheses and are computed using the method proposed in Wu (1986). CP_5 is the prediction from a regression of excess returns on 1-year zero rates and 1-year forward rates at 1-, 2-, 3-, and 4-years. CP_{10} is the prediction from a regression of excess returns on 1-year zero rates and 1-year forward rates at 1-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, and 10-years. $CP_{5,10}$ uses only 5 forward rates as regressors ranging up to 10 years. 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 measuring 1-, 3-, 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 cap prices were measured with error.

On average, the R^2s across different maturities for the $A_1(3)^o$ model are larger than the R^2s for the $A_1(3)$ model by a factor of 1.9. The improvement is even larger for the $A_2(3)^o$ model where the R^2s are larger than the those for the $A_2(3)$ model by an average factor of 2.8. Table 2 contains the difference in R^2's between the $A_1(3)^o$ and $A_1(3)$ models, and between the $A_2(3)^o$ and $A_2(3)$. The bootstrapped confidence intervals indicate that, when we use options to estimate the models, the improvement in R^2 is statistically significant.

Comparing across the models that we estimate with options, the $A_2(3)^o$ model is slightly better than the $A_1(3)^o$ model at predicting 1 year excess returns. Moreover, the R^2s for both models 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 only match excess returns and so they serve as somewhat of an upper bound for the the level of predictability of excess returns.

Figure 1 plots the 1-year expected excess return on a 5-year zero-coupon bond for each of the models that we estimate. Consistent with the results in Tables 1 and 2, 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 $A_0(3)$ 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, 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? In order to facilitate comparisons across models, we examine their implications for the actual and risk-neutral dynamics of the 3-month, 2-year, and 10-year zero-coupon yields.[16] The dynamics of these yields, $R_t = [r_t^{0.25}, r_t^2, r_t^{10}]^\top$, under the actual measure are given by

$$dR_t = \mathcal{K}^{\mathbb{P}} \left(\theta^{\mathbb{P}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 r_t^{0.25} + \Sigma^2 r_t^2 + \Sigma^3 r_t^{10}}\, dW_t\,, \tag{12a}$$

while their dynamics under the risk-neutral measure $\mathbb{Q}$ follow

$$dR_t = \mathcal{K}^{\mathbb{Q}} \left(\theta^{\mathbb{Q}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 r_t^{0.25} + \Sigma^2 r_t^2 + \Sigma^3 r_t^{10}}\, dW_t^{\mathbb{Q}}\,. \tag{12b}$$

[16] Collin-Dufresne et al. (2008) provide a compelling argument for using observables (rather than latent states) to present a standardized comparison across term structure models.

	$A_1(3)^o - A_1(3)$		$A_2(3)^o - A_2(3)$	
	3 Month	1 Year	3 Month	1 Year
2 Yr	2.6	-6.0	3.6	-8.3
	[-11.9, 7.4]	[-12.0, 0.00]	[-1.7, 9.1]	[-15.4, -1.5]
3 Yr	4.6	-0.9	7.0	3.0
	[0.1, 9.4]	[-6.5, 5.1]	[1.6, 12.8]	[-3.0, 9.6]
4 Yr	5.8	2.5	9.0	10.6
	[1.4, 10.7]	[-3.5, 8.6]	[3.8, 14.8]	[4.5, 17.5]
5 Yr	6.5	5.5	10.0	15.6
	[-0.5, 11.8]	[2.1, 11.2]	[5.0, 15.8]	[9.5, 22.5]
6 Yr	6.9	8.2	10.6	18.9
	[2.5, 11.7]	[2.2, 14.9]	[5.4, 16.4]	[12.9, 25.8]
7 Yr	7.2	11.4	10.8	21.4
	[2.7, 12.3]	[5.0, 18.1]	[5.8, 16.7]	[15.3, 28.4]
8 Yr	7.4	13.0	11.0	22.6
	[3.0, 12.4]	[6.6, 20.1]	[5.9, 16.8]	[16.7, 29.7]
9 Yr	7.6	15.1	11.0	23.7
	[3.2, 12.3]	[8.8, 22.2]	[6.2, 16.6]	[17.7, 30.6]
10 Yr	7.5	16.6	10.9	24.2
	[3.2, 12.2]	[9.9, 23.8]	[6.3, 16.4]	[18.4, 31.3]

Table 2: Difference in R^2's

This table presents the difference in R^2's obtained from overlapping weekly projections of one-year zero-coupon swap rate returns, for different maturities, on model-implied 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 method proposed in Wu (1986). We estimated the $A_1(3)$ and $A_2(3)$ models by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 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 cap prices were measured with error.

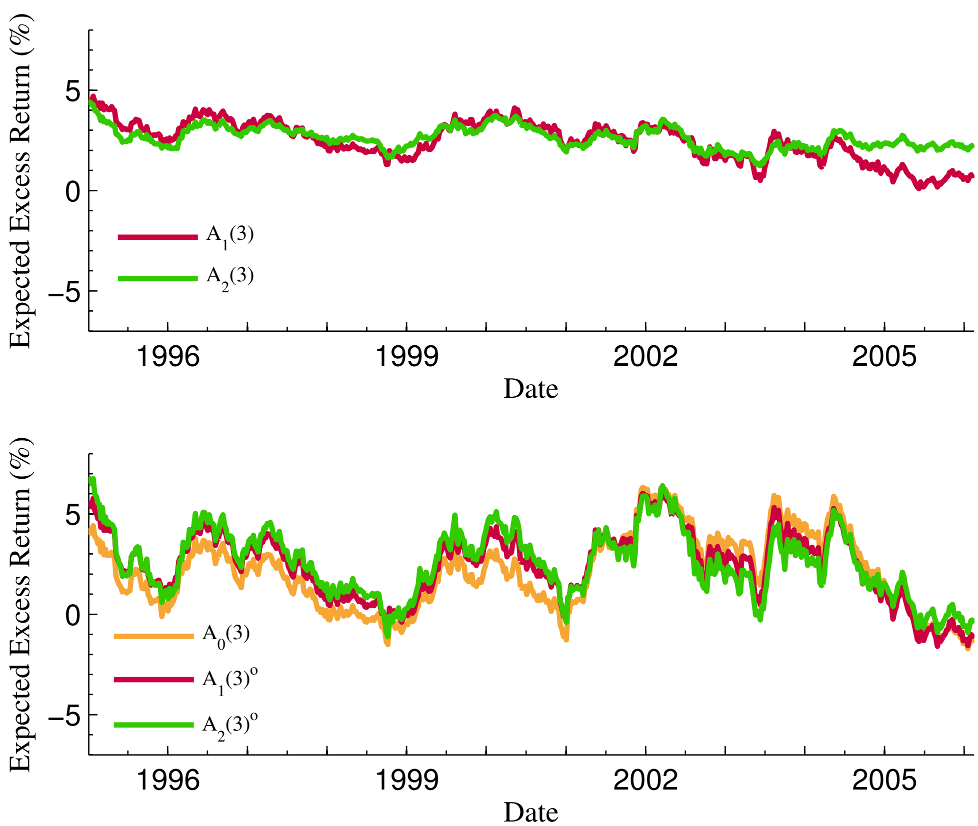


Figure 1: Annual Expected Excess Return – 5-Year Zero-Coupon Bond This figure plots the 1-year expected excess return on a 5-year zero coupon bond for each of the models that we estimate. The top plot shows the expected excess return for the $A_1(3)$ and $A_2(3)$ models that we estimated without using 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.

The estimates of the risk-neutral long-run means, $\theta^{\mathbb{Q}}$, as well as the properties of $\mathcal{K}^{\mathbb{Q}}$ are provided in Table 3. The full parameter estimates are provided in Tables 5, 6, and 7 in Appendix A. Sections B and C in the appendix present the cross-sectional pricing errors and the fit to historical estimates of conditional volatility.

Risk premia in term structure models depend critically on the difference between the actual and risk-neutral distribution of interest rates. The rows of Table 3 that are highlighted in bold illustrate a significant difference in the risk-neutral distribution between the models that are estimated with and without including interest rate options. The risk-neutral long-run means ($\theta^{\mathbb{Q}}_{3\text{m}}$, $\theta^{\mathbb{Q}}_{2\text{y}}$, and $\theta^{\mathbb{Q}}_{10\text{y}}$) 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 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 $A_1(3)^o$ model and 8.46 years in the $A_2(3)^o$ 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 $A_1(3)^o$ and $A_2(3)^o$ 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. We also feel that the lower long-run mean represent a more plausible level of risk aversion with a rate of mean reversion that is on the order of business cycle frequency.

Next we examine how options affect the estimated price of volatility risk. As a measure of the price of this risk, we use the difference between the long run risk-neutral expected zero-coupon bond volatility and the long run actual

Parameter	Estimate				
	$\mathbb{A}_0(3)$	$\mathbb{A}_1(3)$	$\mathbb{A}_1(3)^{\text{o}}$	$\mathbb{A}_2(3)$	$\mathbb{A}_2(3)^{\text{o}}$
$\theta^{\mathbb{Q}}_{3\text{m}}$	**11.1**	**10.5**	**8.02**	**13**	**8.14**
$\theta^{\mathbb{Q}}_{2\text{y}}$	**11.1**	**10.5**	**8.02**	**13**	**8.14**
$\theta^{\mathbb{Q}}_{10\text{y}}$	**11**	**10.4**	**7.86**	**12.7**	**7.98**
half-life$^{\mathbb{Q}}_1$	0.533	0.492	0.436	0.492	0.481
half-life$^{\mathbb{Q}}_2$	1.02	1.12	1.22	1.06	1.16
half-life$^{\mathbb{Q}}_3$	**23.3**	**22.4**	**8.09**	**31.4**	**8.46**
eigenvector$_{3\text{m}}$	0.609	0.609	0.656	0.605	0.653
eigenvector$_{2\text{y}}$	0.593	0.593	0.610	0.593	0.609
eigenvector$_{10\text{y}}$	0.527	0.528	0.445	0.531	0.450

Table 3: Mean Reversion and Long-Run Means

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 1-, 3-, 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 cap prices were measured with error. Each of the models is rotated so that the 3-month, 2-year, and 10-year zero-coupon yields ($R_t = [r_t^{0.25}, r_t^2, r_t^{10}]^\top$) are the factors with drift-adjusted dynamics

$$dR_t = \mathcal{K}^{\mathbb{Q}} \left(\theta^{\mathbb{Q}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1\, r_t^{0.25} + \Sigma^2\, r_t^2 + \Sigma^3\, r_t^{10}}\, dW_t\,. \qquad (12b)$$

under the risk-neutral measure $\mathbb{Q}$. half-life$^{\mathbb{Q}}_i$ is the half-life corresponding to the ith eigenvalue of $\mathcal{K}^{\mathbb{Q}}$. eigenvector$_{\{3\text{m},2\text{y},10\text{y}\}}$ is the third eigenvector of $\mathcal{K}^{\mathbb{Q}}$.

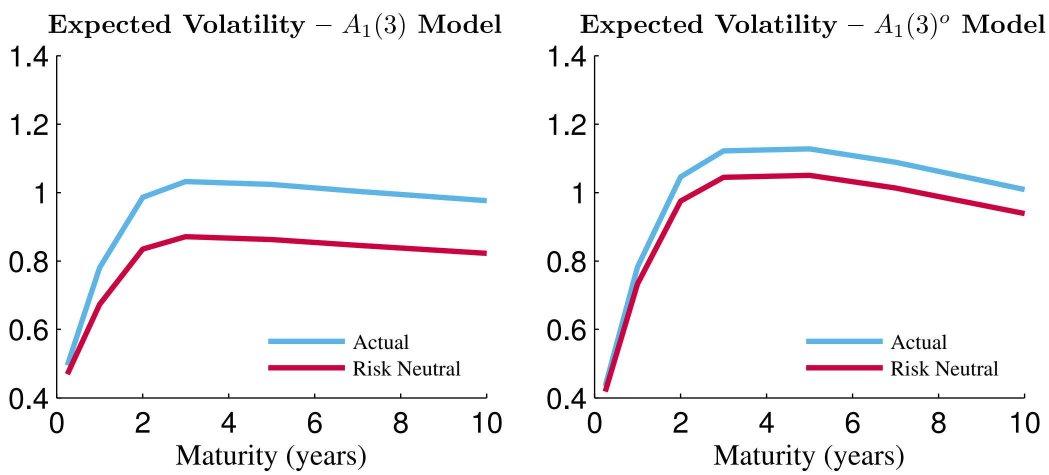


Figure 2: Actual and Risk Neutral Expected Volatility - $A_1(3)$
These figures show the long run expected volatility of zero coupon yields with maturities up to 10 years. The plot on the left shows the expected long run volatility in the $A_1(3)$ model that is estimated without using options. The plot on the right shows the expected long run volatility in the $A_1(3)^o$ model that is estimated using options. The dashed line is the expected volatility under the risk neutral pricing measure. The solid line is the actual expected volatility.

expected zero-coupon bond volatility.[17] Figures 2 and 3 plot the actual long run expected volatility and the risk-neutral long run expected volatility of zero-coupon swap rates for the models with stochastic volatility. This measure of the price of volatility risk is essentially zero for the $A_2(3)^o$ model that we estimated with options and small but negative for the $A_1(3)^o$ model that we estimated with options. When we do not include options in estimation, this measure of the price of volatility risk is more negative in the $A_1(3)$ model, and

[17]The long run expected volatility of the τ-year zero-coupon swap rate is

$$\frac{1}{\tau}\sqrt{B(\tau)^{\top}\Delta\left[\alpha-\beta\left(\mathcal{K}_1^{\mathbb{P}}\right)^{-1}\mathcal{K}_0^{\mathbb{P}}\right]B(\tau)}.$$

The long run risk neutral expected volatility of the τ-year zero-coupon swap rate is

$$\frac{1}{\tau}\sqrt{B(\tau)^{\top}\Delta\left[\alpha-\beta\,\mathcal{K}_1^{-1}\mathcal{K}_0\right]B(\tau)}.$$

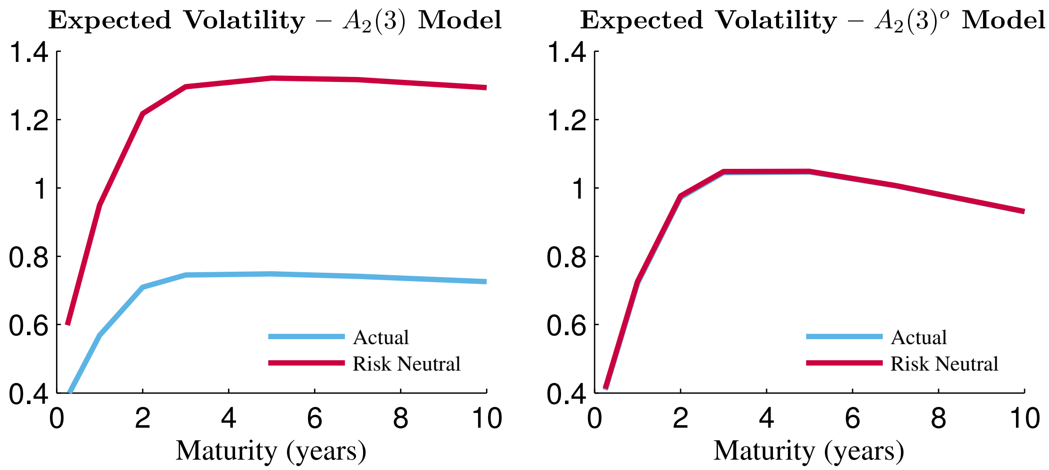


Figure 3: Actual and Risk Neutral Expected Volatility - $A_2(3)$
These figures show the long run expected volatility of zero coupon yields with maturities up to 10 years. The plot on the left shows the expected long run volatility in the $A_2(3)$ model that is estimated without using options. The plot on the right shows the expected long run volatility in the $A_2(3)^o$ model that is estimated using options. The dashed line is the expected volatility under the risk neutral pricing measure. The solid line is the actual expected volatility.

large and positive for the $A_2(3)$ model. Therefore, including options also affects the price of volatility risk that we estimate. Since the $A_1(3)^o$ and $A_2(3)^o$ are best at pricing interest rate caps and predicting excess returns, these models indicate that the price of volatility risk is small, and possibly negative.

4 Linear Projection of Yields

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 LPY(I), is to match the pattern of violations of the expectations hypothesis 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 it is negative (or 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.

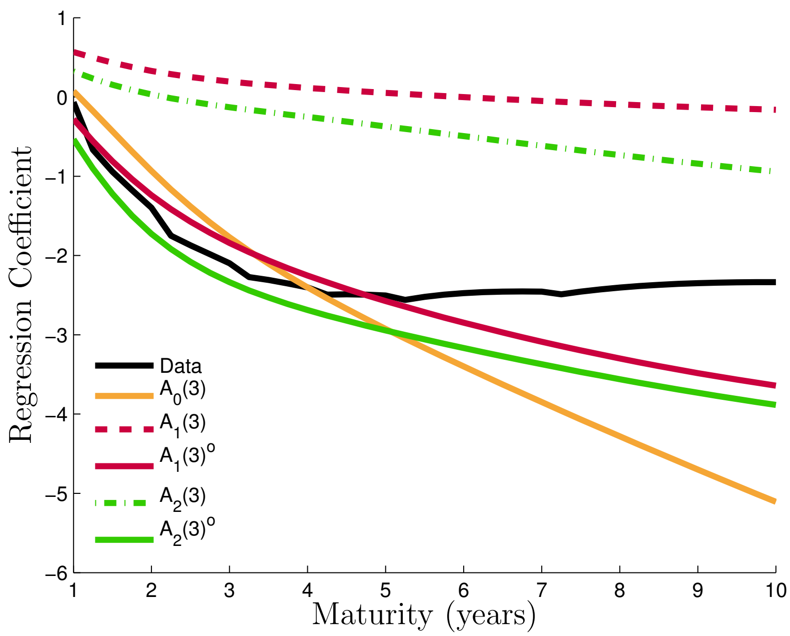


Figure 4: Regression Coefficients from Linear Projection on Yields This figure shows the regression coefficients of the Campbell-Shiller regression $R_{t+1}^{n-1} - R_t^n$ on slope, $(R_t^n - r_t)/(n-1)$. The model values are simulated mean regression coefficients. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 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 cap prices were measured with error.

Specifically, Fama and Bliss (1987) and Campbell and Shiller (1991) perform the following regression

$$r_{t+\Delta t}^{n-\Delta t} - r_t^n = \Phi_{0,n} + \Phi_{1,n} \frac{\Delta t}{n - \Delta t} \left(r_t^n - r_t^{\Delta t} \right) + \varepsilon_t^n ,$$

and find that the regression coefficients, $\hat{\Phi}_{1,n}$, are increasingly negative for larger maturities n. Figure 4 provides the LPY(I) regression results for the models that we estimate.[18,19] Dai and Singleton (2002) find that only the $A_0(3)$ model matches LPY(I), but we find that all of the models are statistically consistent with the observed slope coefficients.[20] 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. Figure 5 shows the 95% simulated confidence interval for the $A_1(3)^o$ model.

The second challenge, LPY(II), states that if the risk premium in the model is correct, then the risk premium adjusted regression

$$\underbrace{r_{t+\Delta t}^{n-\Delta t} - r_t^n + \frac{\Delta t}{n-\Delta t}\mathbb{E}_t\left[r_{t,\Delta t}^{e,n}\right]}_{\text{PACY}_{t,\Delta t}^n} = \Phi_{0,n} + \Phi_{1,n}\underbrace{\frac{\Delta t}{n-\Delta t}\left(r_t^n - r_t^{\Delta t}\right)}_{\text{SLOPE}_t^n} + \varepsilon_t^n\,, \quad (13)$$

should produce a regression coefficient $\hat{\Phi}_{1,n} = 1$. We find that the combination of small sample size and near unit roots in zero-coupon swap rates results in a small, but non-negligible, bias in the regression coefficients. The source of this bias is described in Appendix D. Though cumbersome, the bias can be computed in closed form and will not, in general, be zero. Instead of directly computing the bias, we estimate it by simulating directly from the model and computing the deviation from unity of the simulated LPY(II) coefficients.

Figure 6 shows the model LPY(II) regression results adjusted for the bias. The $A_1(3)^o$ and $A_2(3)^o$ models that we estimate with options dominate the $A_0(3)$, $A_1(3)$, and $A_2(3)$ models that we estimate without options. Although we do not recover an exact regression coefficient of one, the value is quite near the center of the simulated confidence intervals for the $A_0(3)$, $A_1(3)^o$, and $A_2(3)^o$ models. Figure 7 shows the simulated 95% confidence interval for the $A_1(3)^o$ model. The stochastic volatility models that we estimated without

[18] We compute the linear projections using 3 month changes in swap rates rather than the 1 month changes that Dai and Singleton (2002) use. We chose 3 month changes to minimize the effect from bootstrapping the zero coupon yield curve. The results for 1 month changes are qualitatively similar.

[19] The mean regression coefficients for each model were generated using 1,000 simulations.

[20] We use the Cheridito et al. (2007) specification for the market prices of risk which is more flexible than the specification that Dai and Singleton (2002) used.

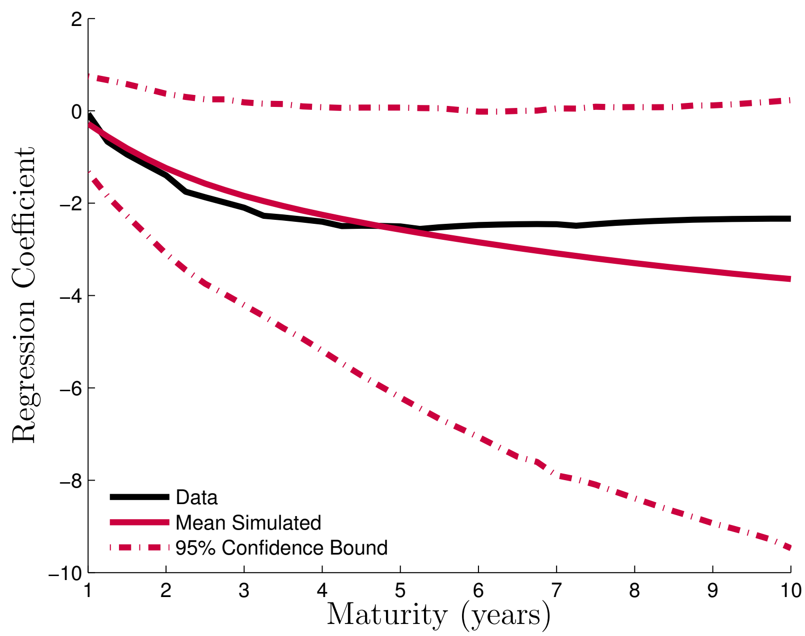


Figure 5: Confidence Interval for Linear Projection on Yields
This figure shows the sample regression coefficient of the Campbell-Shiller regression $R_{t+1}^{n-1} - R_t^n$ on slope, $(R_t^n - r_t)/(n-1)$ for the $A_1(3)^o$ model. The dotted line provides the confidence interval computed from simulation. The model was estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 5-, and 7-year zeros and 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money cap prices with error.

options nearly follow the 95% lower confidence bound. In summary, when we include options in estimation, the expected excess returns in the $A_1(3)^o$ and $A_2(3)^o$ models match the desired properties of linear projections on yields proposed in Dai and Singleton (2003).

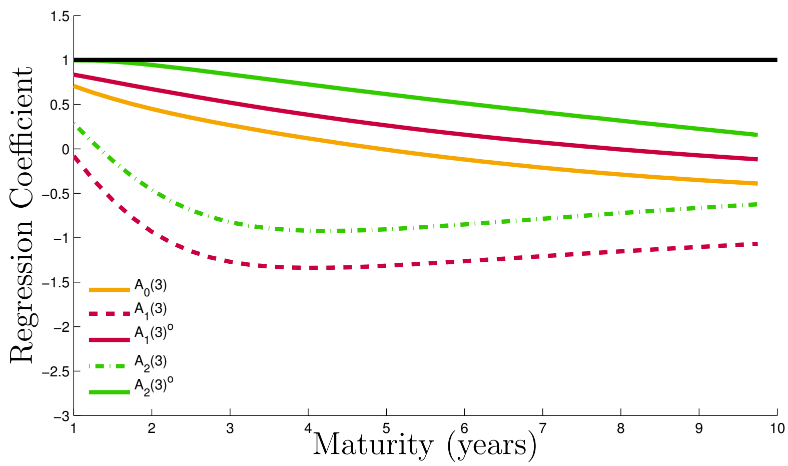


Figure 6: Risk Premium Projection Adjusted for Bias
This figure shows the regression coefficient from the projection of risk premium adjusted excess returns on the slope, $(R_t^n - r_t)/(n-1)$. The regression coefficient is adjusted for small sample bias which we compute via simulation. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 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 cap prices were measured with error.

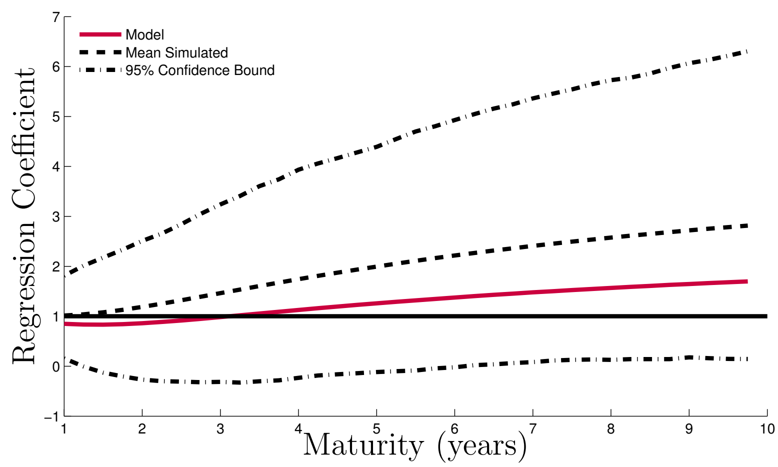


Figure 7: Confidence Interval for LPY(II)
This figure shows the confidence interval for the regression coefficient in the $A_1(3)^o$ model from the projection risk premium adjusted excess returns on the slope, $(R_t^n - r_t)/(n-1)$. The solid line is the regression coefficient for the $A_1(3)^o$. The difference between the solid line at 1 and the dashed line (labeled 'Mean Simulated') represents the small sample bias in the regression coefficient. The lines in Figure 6 are adjusted for this bias.. The $A_1(3)^o$ model is estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 5-, and 7-year zeros and 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money cap prices with error.

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 of model depend on the level, slope, and curvature of yields (which are, in turn, affine functions of the underlying states in the model). In each model that we estimate we can write

$$\begin{aligned} \mathbb{E}_t\left[r^{e,\tau}_{t,\Delta t}\right] &= \texttt{CONSTANT} + \texttt{LEVEL} \times r^{0.5}_t + \texttt{SLOPE} \times \left(r^{10}_t - r^{0.5}_t\right) \\ &\quad + \texttt{CURVATURE} \times \left(2 \cdot r^2_t - r^{10}_t - r^{0.5}_t\right) . \end{aligned} \tag{14}$$

Table 4 provides the values for `CONSTANT`, `LEVEL`, `SLOPE`, and `CURVATURE` for each of the models that we estimate. The value of `CURVATURE` is virtually the same in the $A_1(3)$ and $A_1(3)^o$ models, but the value of `LEVEL` increases slightly from 1.2019 to 1.5896 and the value of `SLOPE` doubles from 1.4992 to 2.9757. Similarly, for the $A_2(3)$ and $A_2(3)^o$ models, the value of `CURVATURE` is very similar, but the value of `LEVEL` increases from 0.4109 to 1.5082 and the value of `SLOPE` increases from 0.6602 to 2.8510. Therefore, 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. For the $A_2(3)^o$ model, including options in estimation also helps to identify the risk premium that is associated with the level of the yield curve. It is interesting to note that the value of `SLOPE` is very similar in the $A_0(3)$, $A_1(3)^o$, and $A_2(3)^o$ models (2.7523, 2.9757, and 2.8510 respectively) which are the best models for predicting excess returns and match the projections of excess returns on the slope of the yield curve.

Comparing across the $A_0(3)$ model and the models that we estimate with options, the $A_0(3)$ and $A_1(3)^o$ models differ mainly in the value of `LEVEL` (1.0793 verus 1.5896 respectively). The $A_0(3)$ and $A_2(3)^o$ models differ in the values of both `LEVEL` (1.0793 verus 1.5082 respectively) and `CURVATURE` (-0.3025 versus 0.2898 respectively). The $A_1(3)^o$ and $A_2(3)^o$ models differ mainly in the value of `CURVATURE` (-0.3561 versus 0.2898 respectively).

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
CONSTANT	-0.0721	-0.0531	-0.0943	0.0001	-0.0840
LEVEL	1.0793	1.2019	1.5896	0.4109	1.5082
SLOPE	2.7523	1.4992	2.9757	0.6602	2.8510
CURVATURE	-0.3025	-0.3517	-0.3561	0.4066	0.2898

Table 4: 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 1-year expected excess return on a 5-year zero-coupon bond is

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

For example, the 1-year expected excess return on a 5-year zero-coupon bond in the $A_1(3)^o$ model is

$$\begin{aligned}\mathbb{E}_t\left[r_{t,1}^{e,5}\right] &= -0.0943 + 1.5896 \times r_t^{0.25} + 2.9757 \times \left(r_t^{10} - r_t^{0.25}\right) \\ &\quad - 0.3561 \times \left(2 \cdot r_t^2 - r_t^{10} - r_t^{0.25}\right) ,\end{aligned}$$

where $r_t^{0.25}$ is the 3-month interest rate, r_t^2 is the 2-year interest rate, and r_t^{10} is the 10-year interest rate.

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 plausible level. We also show that including options reduces our estimate of the magnitude of the price of volatility risk, and helps us to identify the portion of the risk premium that is associated with the slope and level 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.

A Estimation

We use quasi-maximum likelihood to estimate model parameters for $A_0(3)$, $A_1(3)$, and $A_2(3)$ models described by equation (2). We estimate all of the models under the assumption that the model correctly prices 3-month Libor and the 2- and 10-year zero-coupon swap rates exactly and we assume that the remaining zero-coupon swap rates are priced with error. 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.

Using the instruments priced without error and the risk neutral dynamics of X_t, we invert to find the time series of states $\{X_t\}$. Given the states, we then compute the model implied prices of the instruments priced without error. Following Dai and Singleton (2002), we assume that the pricing errors are IID normal with mean zero. Finally, using the physical dynamics of the state vector and the QML approximation, we compute the likelihood of the inverted states. This gives the liklelihod of a given set of parameters to be:

$$\text{likelihood} = \prod \ell^P_{QML}(X_t|X_{t-1}) \cdot (\text{Jacobian}) \cdot (\text{likelihood of pricing errors}).$$

We use a slighlty different procedure than Duffee (2002) to compute the conditional mean and variance of the state variable. For a general affine process, X_t, with conditional drift $K_0 + K_1 X_t$ and conditional variance $H_0 + H_1 \cdot X_t$, the mean and variance of X_t conditional on X_0 satisfy the differential equations

$$\begin{aligned} \dot{M}_t &= K_0 + K_1 M_t\,, \\ \dot{V}_t &= K_1 V_t + V_t K_1^\top + H_0 + H_1 \cdot M_t\,, \end{aligned}$$

If we let f be the $(N + N^2)$-vector $(M, \text{vec}(V))$, then by stacking these coupled ODEs we see that f satisfies the ODE

$$\dot{f} = \begin{bmatrix} K_1 & 0 \\ \Delta & I_N \otimes K_1 + K_1 \otimes I_N \end{bmatrix} f + \begin{bmatrix} K_0 \\ \text{vec}(H_0) \end{bmatrix},$$

where Δ is an $(N^2 \times N)$ matrix with $\Delta_{i,j} = \text{vec}(H_{1,\cdot,\cdot,i})_j$. Rather than considering separate cases to solve this ODE in closed form, we instead compute the fundamental solution numerically using 4th order Runge-Kutta. From the fundamental solution, it is then easy to compute the solution for arbitrary initial conditions.

In each of the models that we estimate, the latent factors are linear combinations of the 3-month, 2-year, and 10-year zero-coupon yields ($r_t^{0.25}$, r_t^2, and r_t^{10} respectively). In order to facilitate comparisons, we characterize each model by its implications for:

- the dynamics of these yields

$$dR_t = \mathcal{K}^{\mathbb{P}} \left(\theta^{\mathbb{P}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 r_t^{0.25} + \Sigma^2 r_t^2 + \Sigma^3 r_t^{10}}\, dW_t \,; \quad (12a)$$

- and the risk-neutral dynamics of these yields

$$dR_t = \mathcal{K}^{\mathbb{Q}} \left(\theta^{\mathbb{Q}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 r_t^{0.25} + \Sigma^2 r_t^2 + \Sigma^3 r_t^{10}}\, dW_t^{\mathbb{Q}} \,. \quad (12b)$$

- the dependence of the short interest rate on these yields,

$$r_t = \rho_0 + [\rho_{3\mathrm{m}}, \rho_{2\mathrm{y}}, \rho_{10\mathrm{y}}] \cdot \underbrace{\left[r_t^{0.25}, r_t^2, r_t^{10}\right]^\top}_{R_t} ; \quad (12c)$$

In equations (12a) and (12b), $\{\Sigma^0, \Sigma^1, \Sigma^2, \Sigma^3\}$ are symmetric positive-definite matrices that govern the covariance of yields. $\theta^{\mathbb{P}}$ is the long-run means of the 3-month, 2-year, and 10-year zero-coupon yields and $\theta^{\mathbb{Q}}$ is their long-run means under the risk-neutral measure. $\mathcal{K}^{\mathbb{P}}$ governs the rate of mean reversion and $\mathcal{K}^{\mathbb{Q}}$ determines the mean reversion under the risk-neutral measure.

Parameter	Estimate				
	$\mathbb{A}_0(3)$	$\mathbb{A}_1(3)$	$\mathbb{A}_1(3)^{\text{o}}$	$\mathbb{A}_2(3)$	$\mathbb{A}_2(3)^{\text{o}}$
ρ_0	-0.00117	-0.00121	-0.00304	-0.0012	-0.00279
$\rho_{3\text{m}}$	1.38	1.39	1.42	1.39	1.4
$\rho_{2\text{y}}$	-0.54	-0.556	-0.605	-0.56	-0.574
$\rho_{10\text{y}}$	0.172	0.178	0.232	0.18	0.216
$\theta^{\mathbb{P}}_{3\text{m}}$	4	3.31	3.86	1.44	4.3
$\theta^{\mathbb{P}}_{2\text{y}}$	4.51	3.96	4.3	1.79	4.58
$\theta^{\mathbb{P}}_{10\text{y}}$	5.5	5.31	5.38	2.82	5.33
$\mathcal{K}^{\mathbb{P}}_{3\text{m},3\text{m}}$	2.07	2.79	2.77	2.41	2.84
$\mathcal{K}^{\mathbb{P}}_{3\text{m},2\text{y}}$	-3.12	-3.99	-4.45	-3.48	-4.32
$\mathcal{K}^{\mathbb{P}}_{3\text{m},10\text{y}}$	1.12	1.55	2.51	1.26	2.18
$\mathcal{K}^{\mathbb{P}}_{2\text{y},3\text{m}}$	-0.362	0.86	-0.185	0.403	-0.869
$\mathcal{K}^{\mathbb{P}}_{2\text{y},2\text{y}}$	0.311	-0.986	0.144	-0.107	1.28
$\mathcal{K}^{\mathbb{P}}_{2\text{y},10\text{y}}$	0.349	0.464	0.468	-0.177	-0.0304
$\mathcal{K}^{\mathbb{P}}_{10\text{y},3\text{m}}$	-1.21	0.00219	-0.971	-0.194	-1.47
$\mathcal{K}^{\mathbb{P}}_{10\text{y},2\text{y}}$	1.21	0.178	1.22	0.531	2.1
$\mathcal{K}^{\mathbb{P}}_{10\text{y},10\text{y}}$	0.371	0.141	0.0189	-0.257	-0.388

Table 5: Parameter Estimates: ρ, $\mathcal{K}^{\mathbb{P}}$, and $\theta^{\mathbb{P}}$.

This table presents the parameter estimates of ρ, $\mathcal{K}^{\mathbb{P}}$, and $\theta^{\mathbb{P}}$ from equations

$$dR_t = \mathcal{K}^{\mathbb{P}} \left(\theta^{\mathbb{P}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 \, r_t^{0.25} + \Sigma^2 \, r_t^2 + \Sigma^3 \, r_t^{10}} \, dW_t \, , \tag{12a}$$

and

$$r_t = \rho_0 + [\rho_{3\text{m}}, \rho_{2\text{y}}, \rho_{10\text{y}}] \cdot \underbrace{\left[r_t^{0.25}, r_t^2, r_t^{10}\right]^\top}_{R_t} , \tag{12c}$$

where $R_t := \left[r_t^{0.25}, r_t^2, r_t^{10}\right]^\top$ is the vector of 3-month, 2-year, and 10-year zero-coupon yields. 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 1-, 3-, 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 cap prices were measured with error.

Parameter	Estimate				
	$\mathbb{A}_0(3)$	$\mathbb{A}_1(3)$	$\mathbb{A}_1(3)^o$	$\mathbb{A}_2(3)$	$\mathbb{A}_2(3)^o$
$\theta^{\mathbb{Q}}_{3m}$	11.1	10.5	8.02	13	8.14
$\theta^{\mathbb{Q}}_{2y}$	11.1	10.5	8.02	13	8.14
$\theta^{\mathbb{Q}}_{10y}$	11	10.4	7.86	12.7	7.98
$\mathcal{K}^{\mathbb{Q}}_{3m,3m}$	2.85	2.92	3.08	2.93	2.95
$\mathcal{K}^{\mathbb{Q}}_{3m,2y}$	-3.97	-4.08	-4.41	-4.1	-4.2
$\mathcal{K}^{\mathbb{Q}}_{3m,10y}$	1.21	1.25	1.63	1.27	1.52
$\mathcal{K}^{\mathbb{Q}}_{2y,3m}$	0.856	0.854	0.853	0.855	0.855
$\mathcal{K}^{\mathbb{Q}}_{2y,2y}$	-0.722	-0.716	-0.709	-0.72	-0.711
$\mathcal{K}^{\mathbb{Q}}_{2y,10y}$	-0.142	-0.146	-0.168	-0.144	-0.168
$\mathcal{K}^{\mathbb{Q}}_{10y,3m}$	0.129	0.119	0.124	0.13	0.141
$\mathcal{K}^{\mathbb{Q}}_{10y,2y}$	-0.0026	0.03	0.0192	-0.00315	-0.00956
$\mathcal{K}^{\mathbb{Q}}_{10y,10y}$	-0.115	-0.14	-0.123	-0.122	-0.11

Table 6: Parameter Estimates: $\mathcal{K}^{\mathbb{Q}}$ and $\theta^{\mathbb{Q}}$.

This table presents the parameter estimates of $\mathcal{K}^{\mathbb{Q}}$ and $\theta^{\mathbb{Q}}$ from

$$dR_t = \mathcal{K}^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - R_t\right) dt + \sqrt{\Sigma^0 + \Sigma^1 r_t^{0.25} + \Sigma^2 r_t^2 + \Sigma^3 r_t^{10}}\, dW_t\,, \qquad (12b)$$

where $R_t := [r_t^{0.25}, r_t^2, r_t^{10}]^\top$ is the vector of 3-month, 2-year, and 10-year zero-coupon yields. 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 1-, 3-, 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 cap prices were measured with error.

Parameter	Estimate				
	$\mathbb{A}_0(3)$	$\mathbb{A}_1(3)$	$\mathbb{A}_1(3)^o$	$\mathbb{A}_2(3)$	$\mathbb{A}_2(3)^o$
$\Sigma^0_{3m,3m}$	0.278	0.174	0.038	0.0173	-0.00649
$\Sigma^0_{3m,2y}$	0.237	0.118	0.0309	0.0217	-0.00685
$\Sigma^0_{3m,10y}$	0.15	0.0741	0.0249	0.0199	-0.00276
$\Sigma^0_{2y,2y}$	0.757	0.19	-0.437	0.0595	-0.212
$\Sigma^0_{2y,10y}$	0.666	0.154	-0.386	0.0999	-0.115
$\Sigma^0_{10y,10y}$	0.853	0.166	-0.445	0.137	-0.109
$\Sigma^1_{3m,3m}$		0.0219	0.0238	0.0725	0.0417
$\Sigma^1_{3m,2y}$		0.0388	0.0217	0.0457	0.0444
$\Sigma^1_{3m,10y}$		0.027	0.0141	0.0116	0.0334
$\Sigma^1_{2y,2y}$		0.242	0.241	0.127	-0.0171
$\Sigma^1_{2y,10y}$		0.204	0.209	0.0221	-0.129
$\Sigma^1_{10y,10y}$		0.244	0.23	-0.00164	-0.181
$\Sigma^2_{3m,3m}$		-0.0658	-0.0809	-0.148	-0.121
$\Sigma^2_{3m,2y}$		-0.117	-0.074	-0.0975	-0.117
$\Sigma^2_{3m,10y}$		-0.0809	-0.0478	-0.0294	-0.0828
$\Sigma^2_{2y,2y}$		-0.726	-0.82	-0.285	-0.419
$\Sigma^2_{2y,10y}$		-0.614	-0.711	-0.0794	-0.183
$\Sigma^2_{10y,10y}$		-0.732	-0.784	-0.0427	-0.134
$\Sigma^3_{3m,3m}$		0.0488	0.0757	0.104	0.103
$\Sigma^3_{3m,2y}$		0.0864	0.0693	0.0801	0.0933
$\Sigma^3_{3m,10y}$		0.06	0.0448	0.0377	0.0635
$\Sigma^3_{2y,2y}$		0.538	0.767	0.273	0.59
$\Sigma^3_{2y,10y}$		0.455	0.665	0.157	0.434
$\Sigma^3_{10y,10y}$		0.542	0.733	0.166	0.444

Table 7: Parameter Estimates: Σ^0, Σ^1, Σ^2, and Σ^3.
This table presents the parameter estimates of Σ^0, Σ^1, Σ^2, and Σ^3 from equation (12). The matrices are symmetric so the lower diagonal elements are not reported in order to conserve space. 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 1-, 3-, 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 cap prices were measured 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 8: Pricing Errors in BPS for Swap Implied Zeros
This table shows the root mean square 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 measuring 1-, 3-, 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 cap prices were measured with error.

B Fit to Yields and Cap Prices

Table 8 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.[21] There is very little difference in the cross-sectional fit between the $A_0(3)$, $A_1(3)$, and $A_2(3)$ models that we estimate without using options. Similarly, there is also 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).

[21] The 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).

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
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 9: Relative Pricing Errors in % for At-the-Money Caps
This table shows the root mean square 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 measuring 1-, 3-, 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 cap prices were measured with error.

Table 9 displays the root mean squared pricing errors in percentage terms for at-the-money caps with various maturities.[22] For all of the models, the percentage pricing errors are worst for 1-year caps and decline as the maturity of the cap increases.[23] 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-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 is 9.2% in the $A_1(3)^o$ model. Similarly,

[22]Figures 8 and 9 provide 3- and 5-year at-the-money cap prices for the $A_1(3)$, $A_1(3)^o$, $A_2(3)$, and $A_2(3)^o$ models. Figure 10 plots the time series of prices in the $A_1(3)^o$ model for at-the-money caps with maturities from 1 to 10 years. The time series of cap prices is similar for the $A_2(3)^o$ model.

[23]The root mean squared relative pricing errors for 1-year caps range from 32.4% to 36.4%. The pricing errors for zero-coupon swap rates are also largest at 1-year. Dai and Singleton (2002) find that a fourth factor is required to capture the short end of the yield curve. We choose to implement more parsimonious 3-factor models because we are primarily interested in predicting changes in long-term interest rates.

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.[24] Driessen et al. (2003) estimate a 3-factor Gaussian HJM model with cap data and report absolute pricing errors (averaged across maturities) of 17.1%. 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 $A_2(3)^o$ that we estimate with caps also provide an excellent fit to the prices of at-the-money swaptions.

[24] Previous 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 (2005) 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.

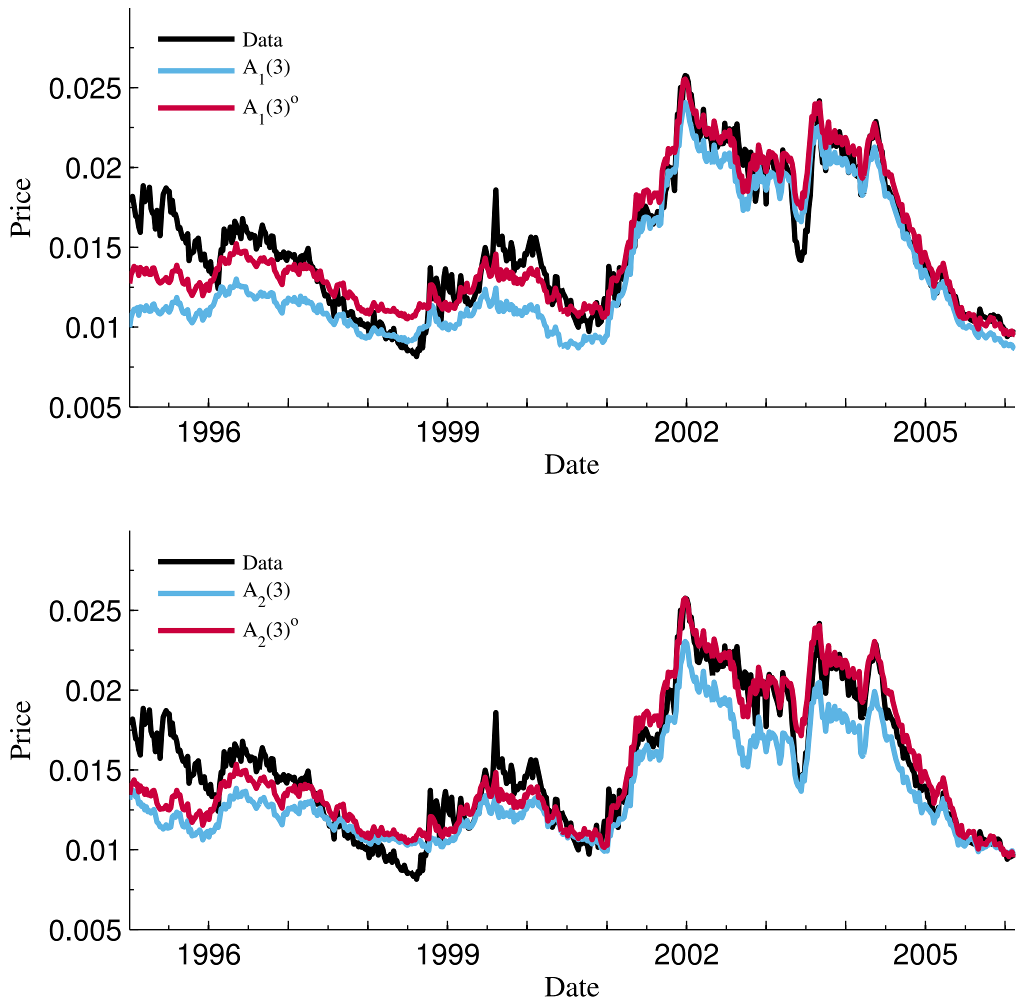
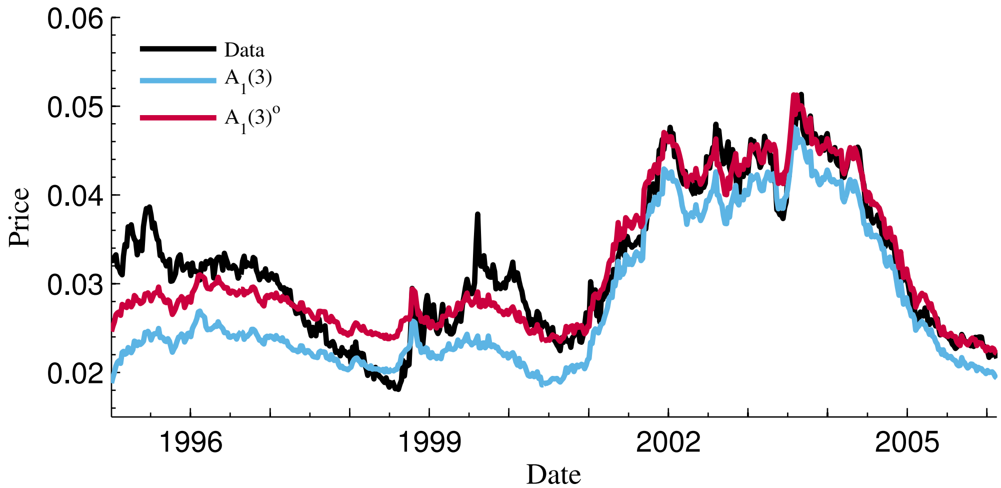


Figure 8: 3-Year At-the-money Cap Prices

The top figure plots the 3-year at-the-money cap prices in the $A_1(3)$ and $A_1(3)^o$ models. The bottom figure plots the 3-year at-the-money cap prices in the $A_2(3)$ and $A_2(3)^o$ models. The $A_1(3)$ and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 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 cap prices were measured with error.



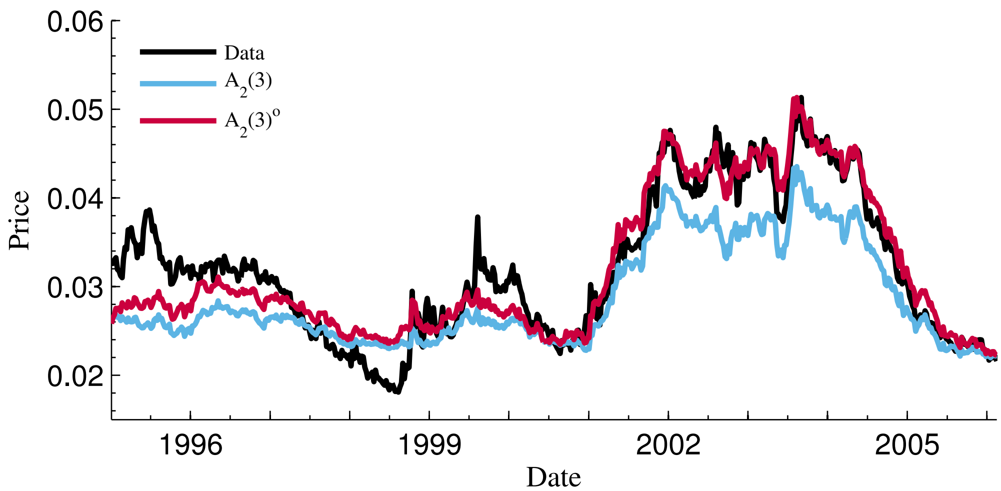


Figure 9: 5-Year At-the-money Cap Prices
The top figure plots the 5-year at-the-money cap prices in the $A_1(3)$ and $A_1(3)^o$ models. The bottom figure plots the 5-year at-the-money cap prices in the $A_2(3)$ and $A_2(3)^o$ models. The $A_1(3)$ and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap zeros and measuring 1-, 3-, 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 cap prices were measured with error.

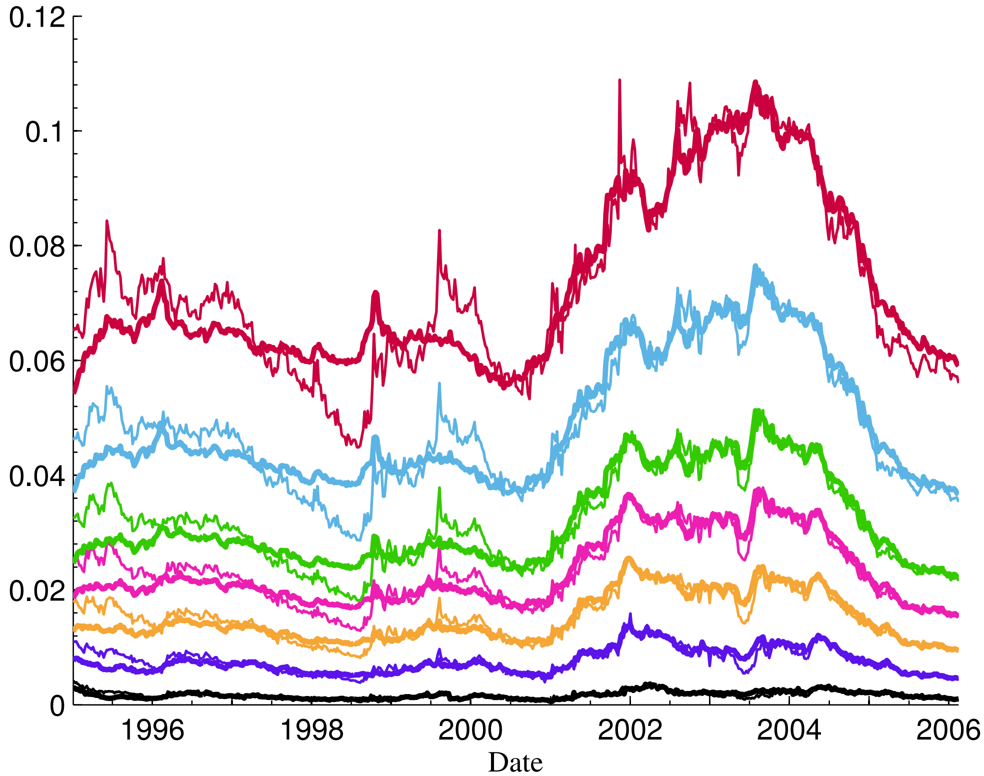


Figure 10: Cap Prices for $A_1(3)^o$ model
This figure plots cap prices for 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps. The think lines indicate actual prices. The model prices for the $A_1(3)^o$ model are plotted in thicker lines, with longer maturities having higher prices. The $A_1(3)^o$ model was estimated by inverting 3-month, 2-year, and 10-year swap zeros. Additionally, the 1-, 3-, 5-, and 7-year zeros and 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money cap prices were measured with error.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
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 10: 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 measuring 1-, 3-, 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 cap prices were measured with error.

C Fit to Volatility

In this section we examine how well the term structure models match the conditional volatility of interest rates. Unlike prices, conditional volatility is not directly observed and therefore it must be estimated.[25] 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. Figures 11, 12, and 13 plot the conditional volatility of zero-coupon swap rates from the term structure models against our estimates of conditional volatility that use historial data. Tables 10 and 11 provide the correlation between the conditional volatilty in the pricing model and the EGARCH and EWMA estimates of conditional volatility respectively. Table 12 provides the average conditional volatilities for the pricing models and the EGARCH and EWMA estimates.

The volatility of the 6-month zero-coupon swap rate is very similar between

[25]Implied 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.

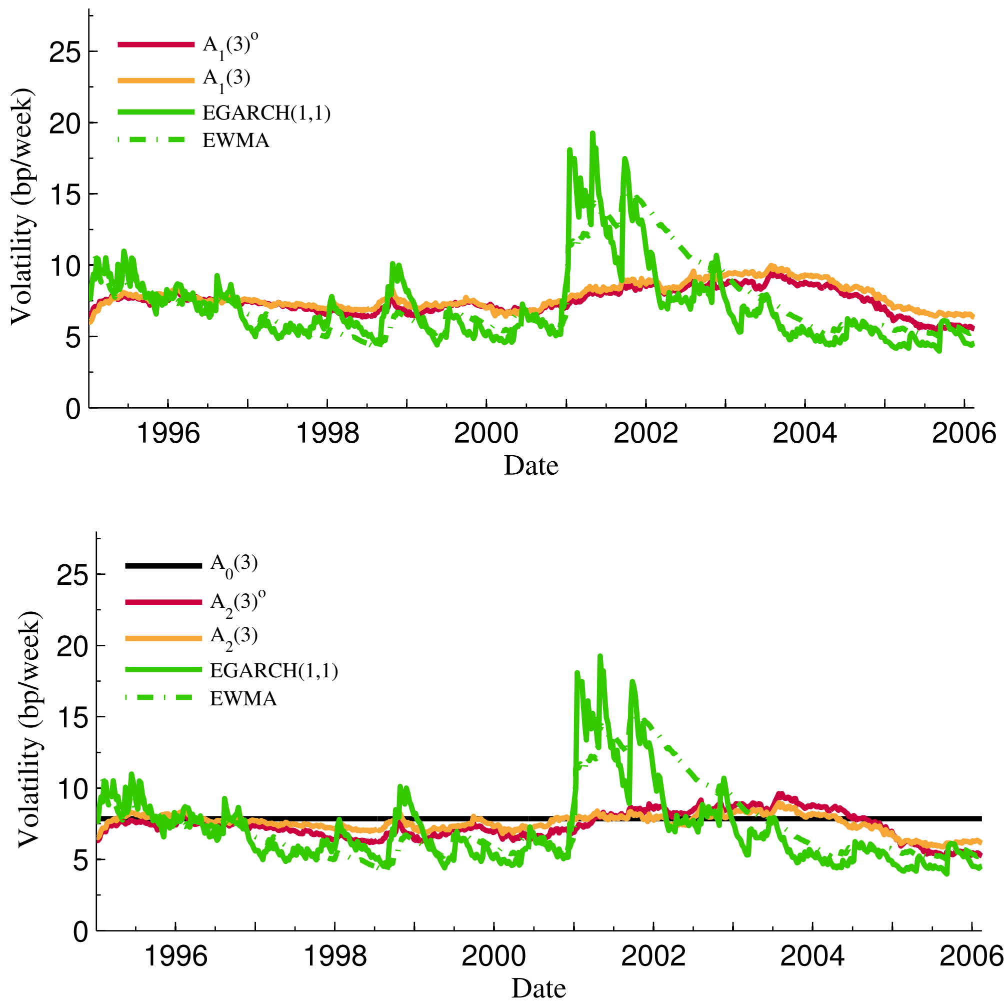


Figure 11: Realized Volatility of 6-Month Zero-Coupon Swap Rate These figures plot weekly model conditional volatility of the 6-month zero coupon swap rate against various estimates of weekly conditional volatility using historical data. For estimates of conditional volatility based on historical data we use an exponential weighted moving average (EWMA) with a 26-week half-life and estimate an EGARCH(1,1) for each maturity. The top plot shows the conditional volatility in the $A_1(3)$ and $A_1(3)^o$ models. The bottom plot shows the conditional volatility in the $A_0(3)$, $A_2(3)$, and $A_2(3)^o$ 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 measuring 1-, 3-, 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 measured with error.

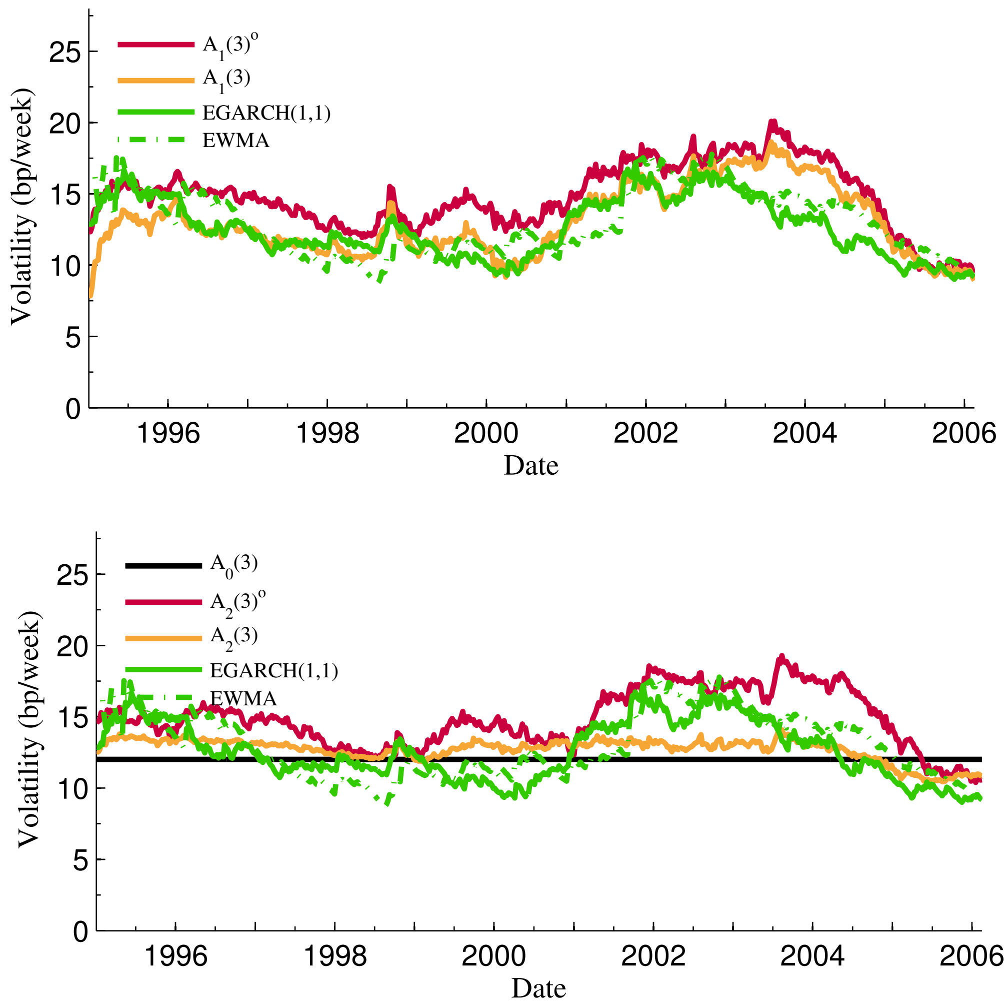
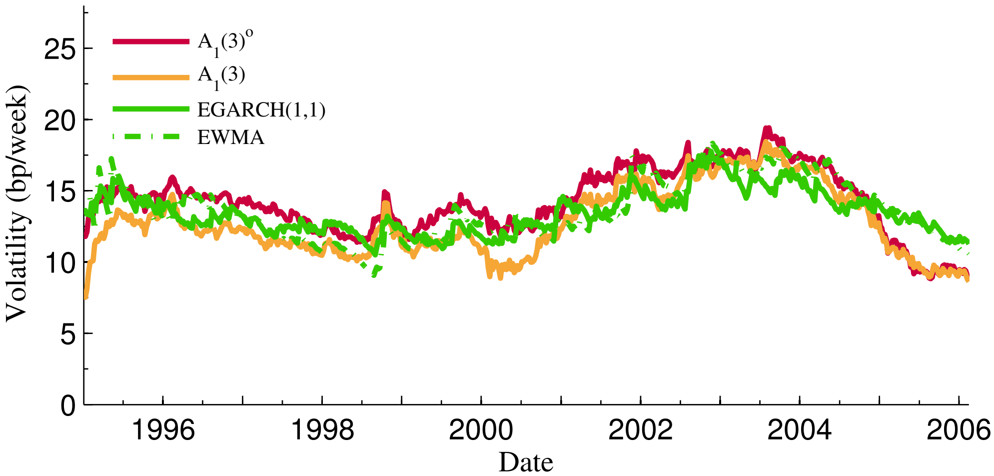


Figure 12: Realized Volatility of 2-Year Zero-Coupon Swap Rate These figures plot weekly model conditional volatility of the 2-year zero coupon swap rate against various estimates of weekly conditional volatility using historical data. For estimates of conditional volatility based on historical data we use an exponential weighted moving average (EWMA) with a 26-week half-life and estimate an EGARCH(1,1) for each maturity. The top plot shows the conditional volatility in the $A_1(3)$ and $A_1(3)^o$ models. The bottom plot shows the conditional volatility in the $A_0(3)$, $A_2(3)$, and $A_2(3)^o$ 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 measuring 1-, 3-, 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 measured with error.



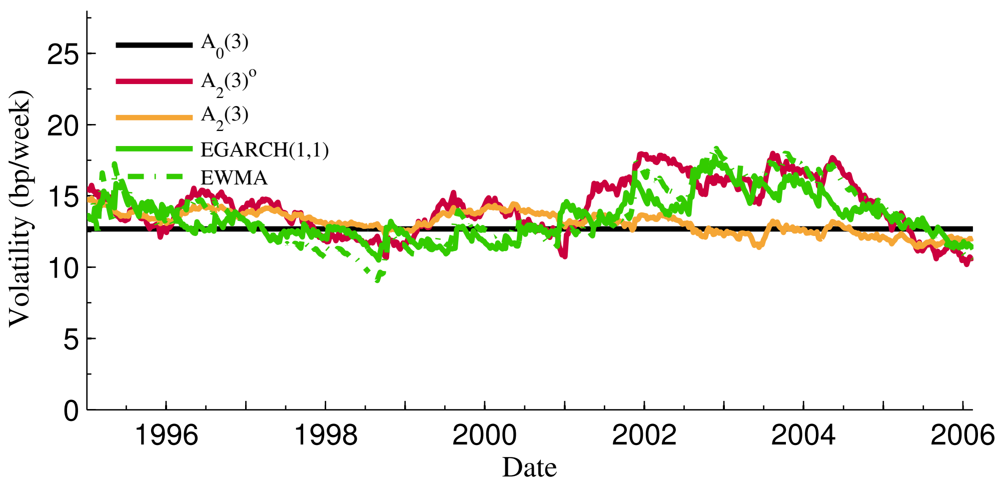


Figure 13: Realized Volatility of 10-Year Zero-Coupon Swap Rate
These figures plot weekly model conditional volatility of the 10-year zero coupon swap rate against various estimates of weekly conditional volatility using historical data. For estimates of conditional volatility based on historical data we use an exponential weighted moving average (EWMA) with a 26-week half-life and estimate an EGARCH(1,1) for each maturity. The top plot shows the conditional volatility in the $A_1(3)$ and $A_1(3)^o$ models. The bottom plot shows the conditional volatility in the $A_0(3)$, $A_2(3)$, and $A_2(3)^o$ 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 measuring 1-, 3-, 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 measured with error.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
6 Month	0	33.2	41.8	35.1	40.3
1 Year	0	55.8	62.7	54.5	63.4
2 Year	0	74.0	76.0	50.1	76.1
3 Year	0	76.7	75.6	29.6	75.6
4 Year	0	77.2	74.1	9.9	74.0
5 Year	0	77.4	73.0	-3.4	72.0
7 Year	0	80.6	74.2	-18.7	71.4
10 Year	0	80.7	75.8	-18.9	71.7

Table 11: Correlation between model and EWMA volatility
This table provides the correlation between model-implied one-week volatilities and Exponential Weighted Moving Average volatility estimates. The EWMA estimates were computing using a 26 week half-life. 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 1-, 3-, 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 measured with error.

the $A_1(3)$ and $A_1(3)^o$ models, and between the $A_2(3)$ and $A_2(3)^o$ models. The historical estimates of the conditional volatility of the 6-month zero-coupon swap rate jump in 2001 and none of the models track this jump.[26] However, all of the pricing models match the average level of the estimates of conditional volatility of the 6-month zero-coupon swap rate.

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 both the EGARCH and EWMA 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 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. The $A_1(3)$ and $A_2(3)$ models match the average level of conditional volatility for

[26]Collin-Dufresne and Goldstein (2002b) find that the volatility of short-term interest rates is not related to the yield curve. Collin-Dufresne et al. (2008) argue that a four-factor model with unspanned stochastic volatility is required to capture the volatility of the short end of the yield curve.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	EGARCH	EWMA
6 Month	7.3	6.8	6.1	6.6	6.1	6.6	6.5
1 Year	9.8	10.5	11.2	10.1	11.1	9.9	10.3
2 Year	12.0	13.1	14.8	12.7	14.8	12.6	12.9
3 Year	12.7	13.6	15.8	13.4	15.9	13.3	13.5
4 Year	12.8	13.6	16.0	13.5	16.1	13.5	13.7
5 Year	12.8	13.5	15.9	13.5	16.0	13.7	13.9
7 Year	12.8	13.2	15.3	13.4	15.4	13.6	13.8
10 Year	12.7	12.8	14.2	13.1	14.2	13.5	13.6

Table 12: Average Conditional Volatilities
This table shows the the average conditional one week volatility of swap rates, $\sigma_t(y_{t+1})$, in basis points. Semi-nonparametric estimates of the the average conditional one week volatility are also provided. 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 1-, 3-, 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 cap prices were measured with error.

the EGARCH and EWMA estimates. The average conditional volatility in the $A_1(3)^o$ and $A_2(3)^o$ models are about 2% higher than the estimates based on EGARCH and EWMA.

The conditional volatility of all swap rates is constant in the $A_0(3)$ model and therefore it 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 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.

D Regression Bias in Yield Projections

We find that the combination of small sample size and near unit roots in zero-coupon swap rates results in a small, but non-negligible, bias in the coefficients for the risk premium adjusted regression in equation (13),

$$\underbrace{r_{t+\Delta t}^{n-\Delta t} - r_t^n + \frac{\Delta t}{n-\Delta t}\mathbb{E}_t\left[r_{t,\Delta t}^{e,n}\right]}_{\text{PACY}_{t,\Delta t}^n} = \Phi_{0,n} + \Phi_{1,n}\underbrace{\frac{\Delta t}{n-\Delta t}\left(r_t^n - r_t^{\Delta t}\right)}_{\text{SLOPE}_t^n} + \varepsilon_t^n .$$

To better understand the source of this bias, note that the regression coefficient in equation (13) is $\hat{\Phi}_{1,n} = U/V$, where

$$\begin{aligned} U &:= \sum_{m=0}^{M-1}\left(\text{PACY}_{m\Delta t,\Delta t}^n - \overline{\text{PACY}^n}\right)\left(\text{SLOPE}_{m\Delta t}^n - \overline{\text{SLOPE}^n}\right) , \\ V &:= \sum_{m=0}^{M-1}\left(\text{SLOPE}_{m\Delta t}^n - \overline{\text{SLOPE}^n}\right)^2 , \end{aligned}$$

and $\overline{\text{PACY}^n}$ and $\overline{\text{SLOPE}^n}$ are the sample averages. In general, $\mathbb{E}[U] \neq \mathbb{E}[V]$.[27] The bias can be approximated by a second order Taylor series expansion:

$$\mathbb{E}\left[\frac{U}{V}\right] = \frac{\mathbb{E}[U]}{\mathbb{E}[V]} + \frac{1}{\mathbb{E}[V]^2}\left(\sigma_V^2 - \text{cov}(U,V)\right) .$$

[27]To see this, note that

$$\mathbb{E}_{m\Delta t}\left[\text{PACY}_{m\Delta t,\Delta t}^n\right] = \text{SLOPE}_{m\Delta t}^n ,$$

therefore

$$\begin{aligned} \mathbb{E}[U-V] &= \sum_{m=0}^{M-1}\mathbb{E}\left[\left(\overline{\text{SLOPE}^n} - \overline{\text{PACY}^n}\right)\text{SLOPE}_{m\Delta t}^n\right] , \\ &= \frac{1}{M}\sum_{m=0}^{M-1}\sum_{k=0}^{m-1}\mathbb{E}\left[\left(\text{SLOPE}_{k\Delta t}^n - \text{PACY}_{k\Delta t,\Delta t}^n\right)\text{SLOPE}_{m\Delta t}^n\right] . \end{aligned}$$

For $k < m$, the residual, $\varepsilon_{k\Delta t} = \mathbb{E}_{k\Delta t}\left[\text{PACY}_{k\Delta t,\Delta t}^n\right] - \text{SLOPE}_{k\Delta t}^n$, may be correlated with $\text{SLOPE}_{m\Delta t}^n$ (and is likely to be more correlated when $\text{SLOPE}_{m\Delta t}^n$ is more nearly stationary). Therefore $\mathbb{E}\left[\text{PACY}_{k\Delta t,\Delta t}^n\text{SLOPE}_{m\Delta t}^n\right] \neq \mathbb{E}\left[\text{SLOPE}_{k\Delta t}^n\text{SLOPE}_{m\Delta t}^n\right]$. Note that this bias is essentially the same as the bias in the regression $y_t = \alpha + \rho y_{t-1} + u_t$ when the true model is $y_t = y_{t-1} + u_t$. See Case 2 in Section 17.4 of Hamilton (1994).

Though cumbersome, the bias can be computed in closed form and is not, in general, zero. Instead of directly computing the bias, we estimate it by simulating directly from the model and computing the deviation from unity of the simulated LPY(II) coefficients.

E Additional Tables and Figures

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

Table 13: Predictability of Excess Returns (R^2s)

This table presents R^2s 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 measuring 1-, 3-, 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 cap prices were measured with error.

References

Dong-Hyun Ahn, Robert F. Dittmar, and A. Ronald Gallant. Quadratic term structure models: Theory and evidence. *The Review of Financial Studies*, 15(1):243–288, Spring 2002.

Yacine Aït-Sahalia and Andrew W. Lo. Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94(1-2):9–51, January-February 2000.

Yacine Aït-Sahalia, Yubo Wang, and Francis Yared. Do option markets correctly price the probabilities of movement of the underlying asset? *Journal of Econometrics*, 102(1):67–110, May 2001.

Torben G. Andersen and Luca Benzoni. Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models. *SSRN eLibrary*, 2008.

Gurdip S. Bakshi, Peter P. Carr, and Liuren Wu. Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. *Journal of Financial Economics*, 87(1):132–156, January 2008.

Ruslan Bikbov and Mikhail Chernov. Term structure and volatility: Lessons from the eurodollar futures and options. Working Paper, 23 November 2005.

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.

Ren-Raw Chen and Louis Scott. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income*, 3:14–31, 1993.

Peng Cheng and Olivier Scaillet. Linear-quadratic jump-diffusion modeling. *Mathematical Finance*, 17(4):575–598, October 2007.

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

Mikhail Chernov and Eric Ghysels. 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(3):407–458, June 2000.

John H. Cochrane and Monika Piazzesi. Bond risk premia. *American Economic Review*, 95(1):138–160, March 2005.

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 interest rate volatility be extracted from the cross section of bond yields? Working Paper, 29 January 2008.

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.

Joost Driessen, Pieter Klaassen, and Bertrand Melenberg. The performance of multi-factor term structure models for pricing and hedging caps and swaptions. *Journal of Financial and Quantitative Analysis*, 38(3):635–672, September 2003.

Jefferson Duarte. Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies*, 17(2):379–404, July 2004.

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

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

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

Bjørn Eraker. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, 59(3):1367–1403, June 2004.

Eugene F. Fama. The information in the term structure. *Journal of Financial Economics*, 13(4):509–528, December 1984a.

Eugene F. Fama. Term premiums in bond returns. *Journal of Financial Economics*, 13(4):529–546, December 1984b.

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

Jeremy J. Graveline. Exchange rate volatility and the forward premium anomaly. Working Paper, 31 January 2008.

James D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

Bin Han. Stochastic volatilities and correlations of bond yields. *Journal of Finance*, 62(3):1491–1524, June 2007.

Steve L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

Jens Carten Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies*, 13(2):433–451, Summer 2000.

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

Christopher S. Jones. The dynamics of stochastic volatility: Evidence from underlying and options markets. *Journal of Econometrics*, 116(1-2):181–224, September-October 2003.

Scott Joslin. Pricing and hedging volatility risk in fixed income markets. Working Paper, 30 September 2007.

Don H. Kim. Spanned stochastic volatility: A reexamination of the relative pricing between bonds and bond options. BIS Working Papers, December 2007.

Markus Leippold and Liuren Wu. Asset pricing under the quadratic class. *The Journal of Financial and Quantitative Analysis*, 37(2):271–295, June 2002.

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

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

Francis A. Longstaff, Pedro Santa-Clara, and Eduardo S. Schwartz. The relative valuation of caps and swaptions: Theory and empirical evidence. *Journal of Finance*, 56(6):2067–2109, December 2001.

Jun Pan. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50, January 2002.

Samuel Brodsky Thompson. Identifying term structure volatility from the libor-swap curve. *The Review of Financial Studies*, 21(2):819–854, April 2008.

Len Umantsev. *Econometric Analysis of European Libor-Based Options within Affine Term-Structure Models*. PhD thesis, Stanford University, 2002.

C.F.J. Wu. Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, 14(4):1261–1295, December 1986.