Do Options Contain Information About Excess Bond Returns?*

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

First Draft: June 3, 2005
Current Draft: March 24, 2006

COMMENTS WELCOME

PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

There is strong empirical evidence that long-term interest rates contain a time-varying risk premium. Interest rate options may contain information about this risk premium because their prices are sensitive to the volatility and market prices of the risk factors that drive interest rates. We use the time series of swap rates and interest rate cap prices to estimate dynamic term structure models. The risk premiums that are estimated using option prices are better able to predict excess returns for long-term swaps over short-term swaps, both in- and out-of-sample. In contrast to previous literature, the most succesful models for predicting excess returns have risk factors with stochastic volatility. We also show that the models that are estimated using option prices are consistent with the failure of the expectations hypothesis.

*We thank Chris Armstrong, Snehal Banerjee, David Bolder, Darrell Duffie, Yaniv Konchitchki, and seminar participants at Stanford GSB graduate student seminar for helpful comments. We are very grateful to Ken Singleton for many discussions and comments.

[†]Ibmec Business School, calmeida@ibmecrj.br

[‡]Stanford Graduate School of Business, jjgravel@stanford.edu

[§]Stanford Graduate School of Business, joslin@stanford.edu

Introduction

A bond or swap that is sold before it matures has an uncertain return. Regressions on the yield curve show that excess returns for long-term bonds and swaps are both predictable and time-varying.[1] This empirical evidence suggests that long-term interest rates contain a time-varying risk premium. Interest rate options may contain information about this risk premium because their prices are sensitive to the volatility and market prices of the risk factors that drive interest rates. We ask whether risk premiums estimated using interest rate option prices are better able to predict changes in long-term interest rates.

We estimate 3-factor affine term structure models using the joint time series of swap rates and interest rate cap prices with different maturities. We then examine how well the estimated risk premiums predict excess returns for long-term swaps over short-term swaps. Previous papers that address this question use only the time-series of bond yields or swap rates to estimate their models. Our main finding is that the risk premiums estimated using interest rate option prices are better able to predict excess returns, both in- and out-of-sample. In short, interest rate option prices contain valuable information about the risk premium in long-term interest rates.

When interest rate options are included in estimation, the models with risk factors that have stochastic volatility are best at predicting excess returns. This result differs from Duffee (2002) and Cheridito et al. (2006) who both find that the predictability of excess returns is best captured by a model with risk factors that have constant volatility. We also find that, with a correction for small sample bias, the models with stochastic volatility successfully capture the failure of the expectations hypothesis and match regressions of returns on the slope of the yield curve. Dai and Singleton (2002) find that only models with constant volatility are successful on this dimension. An obvious additional benefit of models with stochastic volatility is that they better capture the variation in interest rate volatility.

Previous papers that have used interest rate options to estimate dynamic term structure models have focused on accurately pricing both options and underlying interest rates. Umantsev (2002) estimates affine models jointly on both swaps and swaptions and analyzes the volatility structure of these

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

markets as well as factors influencing the behavior of interest rate risk premia. Longstaff et al. (2001) and Han (2004) explore the correlation structure in yields that is required to simultaneously price both caps and swaptions. Bikbov and Chernov (2004) use both Eurodollar futures and short-dated option prices to estimate affine term structure models and discriminate between various volatility specifications. Our paper differs from these papers in that we examine how including options in estimation impacts a model's ability to capture the dynamics of interest rates and predict excess returns.

The remainder of the paper is organized as follows. Section 1 discusses excess returns in fixed income markets. Section 2 describes the data and estimation procedure we use. Section 3 presents the fit to the cross-section of swap rates and cap prices with different maturities. Section 4 examines the fit to historical estimates of conditional volatility. Section 5 compares the estimated models' ability to predict excess returns and Section 6 examines linear projections of excess returns on yields. Section 7 concludes. Technical details, and all tables and figures are provided in appendices

1 Excess Returns in Fixed Income Markets

A bond or swap that is sold before it matures has an uncertain return. For example, although the 5-year interest rate is known today, the return on a 5-year bond that is sold in one year is uncertain. Interest rate volatility is one measure of the amount of such risk that a bond is exposed to. Investors may demand a premium for holding this risk. In this section we use regression analysis to test whether interest rate volatility explains variations in swap returns.

To fix notation, let $P_t^{t+\tau}$ be the price at time t of a zero coupon bond that matures τ years later. Let r_t^τ be the continuously compounded yield on that bond so that

$$r_t^\tau := -\frac{1}{\tau}\ln\left(P_t^{t+\tau}\right) .$$

In Δt years, the new price of the bond is $P_{t+\Delta t}^{\tau-\Delta t}$. Let $r_{t,\Delta t}^{e,\tau}$ be the continuously compounded return on the bond over this period, in excess of the riskless return $r_t^{\Delta t}$. That is,

$$
\begin{aligned}
r_{t,\Delta t}^{e,\tau} &:= \frac{1}{\Delta t}\left[\ln\left(P_{t+\Delta t}^{t+\tau}/P_t^{t+\tau}\right) + \ln P_t^{t+\Delta t}\right] , \\
&= \frac{1}{\Delta t}\left[-\left(\tau - \Delta t\right)\left(r_{t+\Delta t}^{\tau-\Delta t} - r_t^{\tau}\right) + \Delta t\left(r_t^{\tau} - r_t^{\Delta t}\right)\right] . \qquad (1)
\end{aligned}
$$

The expectations hypothesis suggests that the expected excess return on a long-term zero coupon bond or swap is constant. Rearranging equation (1), the expectations hypothesis implies that

$$
\mathbb{E}_t\left[r_{t+\Delta t}^{\tau-\Delta t} - r_t^{\tau}\right] = \text{constant} + \frac{\Delta t}{\tau - \Delta t}\left(r_t^{\tau} - r_t^{\Delta t}\right) . \qquad (2)
$$

If the expectation hypothesis holds, then regressions based on equation (2) should yield a regression coefficient of 1. However, Fama and Bliss (1987) and others find that the expectations hypothesis fails and the regression coefficients are actually negative with an absolute value that increases with the maturity of the bond under consideration.

If investors demand a time-varying premium for holding long term bonds with an uncertain return, then interest rate volatility may provide additional predictive power in a regression. In the subsequent regression analysis, we use the implied volatility from at-the-money interest rate cap prices which provides a forward looking measure of volatility that also incorporates risk preferences.[2]

As a preliminary test of this hypothesis, we regress the one year excess returns of the n-year zero coupon swap (where $n = 2, 3, 4, 5$) on three sets of explanatory variables (all include a constant):

1. the slope of the yield curve, taken as $r_t^n - r_t^1$,

2. the slope and n-year interest rate cap implied volatility,

3. 1- to 5-year zero rates.

We report the R^2 from the regressions using 581 weekly observations from June 1995 to Februrary 2006 in Table 1. The results indicate that including the cap implied volatility in the regression increases the amount of variation

[2] An interest rate cap is a financial derivative that caps the interest rate that is paid on the floating side of a swap. The market convention is to quote prices in terms of the volatility implied by Black's formula.

in excess returns which is explained. However, it should be noted that the sample size is relatively small and the regressions choose coefficients to maximize the R^2 by construction (in particular there are only 10 non-overlapping one year returns.)

The preliminary evidence in these regressions indicates that excess swap returns depend on interest rate volatility, and suggests that it may be beneficial to incorporate interest rate cap prices into a dynamic model of the term structure of interest rates. We turn to this objective in the next section.

2 Model and Estimation Strategy

We estimate three 3-factor affine term structure models with 0, 1, or 2 factors having stochastic volatility.[3] The pricing kernel follows a diffusion process of the form

$$dM_t = -M_t \, r_t \, dt - M_t \, \Lambda_t^\top d\, W_t \,,$$

where

$$\begin{aligned} r_t &:= \rho_0 + \rho_1 \cdot X_t \,, \\ \Lambda_t &:= \left(\sqrt{\Delta \left[\alpha + \beta X_t \right]} \right)^{-1} \left[\left(\mathcal{K}_0^{\mathbb{P}} - \mathcal{K}_0 \right) + \left(\mathcal{K}_1^{\mathbb{P}} - \mathcal{K}_1 \right) X_t \right] \,,^4 \end{aligned}$$

and

$$dX_t = \left[\mathcal{K}_0^{\mathbb{P}} + \mathcal{K}_1^{\mathbb{P}} X_t \right] dt + \sqrt{\Delta \left[\alpha + \beta X_t \right]} \; d\, W_t \,.^5$$

We have used the notation $\Delta\left[\cdot\right]$ to denote a square matrix with its vector argument along the diagonal. The drift, r, of the pricing kernel is commonly referred to as the short interest rate and the volatility, Λ, is commonly referred to as the market price of risk.

---

[3]See Dai and Singleton (2000) for a detailed specification of the $A_M(N)$ affine term structure models that we estimate in this paper.

[5]We use an extended affine market price of risk introduced by Cheridito et al. (2006) as a generalization of the essentially affine market price of risk used in Duffee (2002). The model specifications are described in more detail in the appendix.

[5]The dynamics of the state vector X under the martingale pricing measure $\mathbb{Q}$ defined by

$$\left. \frac{d\,\mathbb{Q}}{d\,\mathbb{P}} \right|_t := e^{\int_0^t r_u \, du} \, \frac{M_t}{M_0} = e^{-\frac{1}{2}\int_0^t \Lambda_u^\top \Lambda_u \, du - \int_0^t \Lambda_u^\top d\, W_u} \,,$$

In this setting, Duffie and Kan (1996) show that the price P_t^T at time t of a zero coupon bond that pays \$1 at time $T \geq t$ is

$$P_t^T = \mathbb{E}_t \left[\frac{M_T}{M_t} 1 \right] = e^{A(T-t)+B(T-t)\cdot X_t} ,$$

where the functions A and B satisfy the Riccati ODEs

$$\begin{aligned} \frac{d}{d\tau} B(\tau) &= -\rho_1 + \mathcal{K}_1^\top B(\tau) + \frac{1}{2} \beta^\top \Delta [B(\tau)] B(\tau) , \quad B(0) = 0 , \\ \frac{d}{d\tau} A(\tau) &= -\rho_0 + \mathcal{K}_0^\top B(\tau) + \frac{1}{2} \alpha^\top \Delta [B(\tau)] B(\tau) , \quad A(0) = 0 . \end{aligned}$$

Using Itô's formula, the dynamics of zero coupon bond prices are given by

$$\begin{aligned} dP_t^T &= P_t^T \left[r_t + B(T-t)^\top \sqrt{\Delta [\alpha + \beta X_t]} \, \Lambda_t \right] dt \\ &\quad + P_t^T B(T-t)^\top \sqrt{\Delta [\alpha + \beta X_t]} \, dW_t . \end{aligned} \tag{3}$$

From (3), the instantaneous expected excess return, or risk premium, at time t for a zero coupon bond that matures at time T is given by

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

This risk premium depends on the volatility, $\sqrt{\Delta [\alpha + \beta X_t]}$, of the risk factors, X_t, and the market prices, Λ_t, of those risk factors.

The results of the regression analysis in the previous section suggest that interest rate options may contain information about the risk premium in interest rates. As such, we also include the prices of interest rate caps in our model estimation. An interest rate cap is a portfolio of options on 3-month Libor that effectively caps the interest rate that is paid on the floating side

---

are

$$dX_t = [\mathcal{K}_0 + \mathcal{K}_1 X_t] \, dt + \sqrt{\Delta [\alpha + \beta X_t]} \, dW_t^{\mathbb{Q}} ,$$

where

$$W_t^{\mathbb{Q}} := W_t + \int_0^t \Lambda_u \, du ,$$

is a Brownian motion under $\mathbb{Q}$.

of a swap. The price of a single option (caplet) on the 3-month Libor interest rate $L^{0.25}_{T-0.25}$ at time $T-0.25$ with strike $\overline{C}$ is

$$\mathbb{E}_t\left[\frac{M_T}{M_t}\,0.25\,\underbrace{\left(L^{0.25}_{T-0.25}-\overline{C}\right)^+}_{\text{caplet payoff}}\right].^6 \tag{4}$$

As equation (4) illustrates, the caplet payoff is sensitive to the volatility of the 3-month interest rate. The price of the caplet is also sensitive to the market prices of risk, Λ, which are embedded in the pricing kernel, M.

The price $C^N_t\left(\overline{C}\right)$ of an N-period interest rate cap with strike rate $\overline{C}$ and 3-month floating interest payments is

$$\begin{aligned} C^N_t\left(\overline{C}\right) &= \sum_{n=2}^{N}\mathbb{E}_t\left[\frac{M_{t+0.25\,n}}{M_t}\,0.25\left(L^{0.25}_{t+0.25(n-1)}-\overline{C}\right)^+\right],^7 \\ &= \sum_{n=2}^{N}\mathbb{E}_t\left[\frac{M_{t+0.25\,n}}{M_t}\left(\frac{1}{P^{t+0.25\,n}_{t+0.25(n-1)}}-\left(1+0.25\,\overline{C}\right)\right)^+\right]. \end{aligned}$$

In the setting of affine term structure models, Duffie et al. (2000) show that cap prices can be computed as a sum of inverted Fourier transforms. However, when the solutions A and B to the Riccati ODEs are not known in closed form, direct Fourier inversion is too computationally expensive for use in estimation. Instead, we use a more computationally efficient adaptive quadrature method that is based on Joslin (2005).

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- and 10-years.[8] Finally, we 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.[9] The full model specifications and estimation pro-

[6]The 3-month Libor interest rate $L^{0.25}_{T-0.25}$ at time $T-0.25$ satisfies

$$1+0.25\,L^{0.25}_{T-0.25}=\frac{1}{P^{T}_{T-0.25}},$$

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

[9]An $A_M(3)$ model has three latent factors with M factors having stochastic volatility. See Dai and Singleton (2000) for more details.

cedure are described in detail in the appendix. All of the models are estimated using the assumption that the model correctly prices 3-month Libor and the 2- and 10-year swap zero coupon rates exactly and the remaining swap zero coupon rates are assumed to be priced with error.[10] For the $A_1(3)^o$ and $A_2(3)^o$ models, we also assume that at-the-money caps with maturities of 1-, 2-, 3-, 4-, 5-, 7-, and 10-years are priced with error. For each model, we use the following procedure to obtain quasi-maximum likelihood estimates:

1. Randomly generate 25 feasible sets of starting parameters.
2. Starting from the best of the feasible seeds, use a gradient search method to obtain a (local) maximum of the quasi-likelihood function constructed using the model's exact conditional mean and variance.[11]
3. Repeat these steps 1000 times to obtain a global maximum.

The parameter estimates are provided in Table 2.

3 Fit to Yields and Cap Prices

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 swap zero rates because the latent states variables are chosen so that the models correctly price these instruments. The root mean squared pricing errors for other maturities range from about 4 basis points to about 10 basis points. There is very little difference in the cross-sectional fit between the $A_0(3)$, $A_1(3)$, and $A_2(3)$ models that are all estimated without using options. Similarly, there is also little difference between the $A_1(3)^o$ and $A_2(3)^o$ models that are both estimated with options. The use of options to estimate the $A_1(3)^o$ and $A_2(3)^o$ models has a only small affect on the models' fit to the cross-section of swap zero coupon rates with different maturities. Including options improves the fit by less than a basis point at the short end of the yield curve (up to 1 year) and worsens the fit by

[10]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. See Chen and Scott (1993) for more details.

[11]In an affine model, the conditional mean and variance are known in closed form as the solution to a linear constant coeffiecient ODE. See Appendix B for details.

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 in percentage terms for at-the-money caps with various maturities.[12] For all of the models, the percentage pricing errors are worst for 1-year caps and decline as the maturity of the cap increases.[13] Amongst the models that are estimated 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 are estimated without options. However, the relative pricing errors for caps with longer maturities are considerably lower when the caps are included 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 $A_1(3)^o$.

To summarize, when options are included in estimation, the fit to zero coupon swap rates with different maturities does not change much, but the fit to at-the-money cap prices improves for longer dated caps.[14]

4 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 must be estimated.[15] For estimates of conditional

[12]Figures 1 and 2 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 3 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.

[13]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 yields.

[14]Although not reported here, the $A_1(3)^o$ and $A_2(3)^o$ that are estimated with caps also provide an excellent fit to the prices of at-the-money swaptions.

[15]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 case

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 4 plots the conditional volatility of zero coupon swap rates in the term structure models against our estimates of conditional volatility that use historial data. Tables 5 and 6 provide the correlation between the conditional volatilty in the pricing model and the EGARCH and EWMA estimates of conditional volatility respectively. Table 7 provides the average conditional volatilities for the pricing models and the EGARCH and EWMA estimates.

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.

The volatility of the 6-month zero coupon swap rate 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. 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. 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 for maturities beyond 4 years. The $A_1(3)$ and $A_2(3)$ models match the average level of conditional volatility for 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.

There are two possible explanations for the difference in the average level of conditional volatility between the models that are estimated with options versus those that are estimated without options. First, option prices could help to better identify the true conditional volatility, which may be higher than our estimates that are based in historical swap rates. The other possi-

the implied volatilities from cap prices to differ from the conditional volatility.

bility is that the models are not flexible enough to match both option prices and the conditional volatility of swap rates. When the $A_1(3)$ model is estimated without including options, it matches our estimates of the conditional volatility of swap rates (which are based on historial swap rate data), but it has larger pricing errors for options. By contrast, when the $A_1(3)^o$ and $A_2(3)^o$ models are estimated using options, they better match option prices, but appear to overstate the conditional volatility of swap rates. Put differently, the $A_1(3)^o$ and $A_2(3)^o$ models better matches implied volatility from interest rate cap prices, but the $A_1(3)$ model better matches our estimates of conditional volatility. In term structure models with stochastic volatility, option-implied volatility and conditional volatility are linked by the market price of volatility risk. The second explanation suggests that the combination of factor volatilities and market prices of risk may not be flexible enough to capture both conditional volatility and option prices. Perhaps a more flexible market price of risk is required, or perhaps a different source of variation, such as jumps, is required.

5 Predictability of Excess Returns

In this section we examine how well the models capture excess returns, or changes in long-term interest rates relative to short-term interest rates.

In an affine term structure model, the expected excess return is given by

$$\begin{aligned}\mathbb{E}_t\left[r^{e,\tau}_{t,\Delta t}\right] &:= \mathbb{E}_t\left[\frac{1}{\Delta t}\ln\left(P^{t+\tau}_{t+\Delta t}/P^{t+\tau}_t\right)+\frac{1}{\Delta t}\ln P^{t+\Delta t}_t\right], \\ &= \frac{1}{\Delta t}\left\{\begin{array}{l} A(\tau-\Delta t)+B(\tau-\Delta t)\cdot\mathbb{E}_t\left[X_{t+\Delta t}\right] \\ -\left[A(\tau)+B(\tau)\cdot X_t\right]+A(\Delta t)+B(\Delta t)\cdot X_t\end{array}\right\}.\end{aligned}$$ [16]

To measure how well each model predicts excess returns, we compute the following statistic

$$R^2 = 1-\frac{\operatorname{var}\left(\mathbb{E}_t\left[r^{e,\tau}_{t,\Delta t}\right]-r^{e,\tau}_{t,\Delta t}\right)}{\operatorname{var}\left(r^{e,\tau}_{t,\Delta t}\right)}.$$

[16] Recall that in an affine term structure model, the price $P^{t+\tau}_t$ at time t of a zero coupon swap that matures at time $t+\tau$ is given by

$$P^{t+\tau}_t = e^{A(\tau)+B(\tau)\cdot X_t}.$$

We then compare the R^2s for each model we estimate with the R^2s from three versions of the regressions of excess returns on forward rates as performed in Cochrane and Piazzesi (2005).[17]

Table 8 presents the R^2 statistics for 3-month excess returns for the period from January 1995 to February 2006 that was used to estimate the model. Amongst the three models that are estimated without options, the $A_1(3)$ model is best at predicting in-sample 3-month 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 in-sample 3-month excess returns, followed by the $A_1(3)$ model and then the $A_2(3)$ model.

When options are included in estimation, the $A_1(3)^o$ and $A_2(3)^o$ modeols are better able to predict in-sample 3-month excess returns relative to their $A_1(3)$ and $A_2(3)$ counterparts that are estimated without including options. 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 1.3. The improvement is even larger for the $A_2(3)$ model where the R^2s are larger than the those for the $A_2(3)$ model by an average factor of 2. The $A_2(3)^o$ model is slightly better than the $A_1(3)^o$ model at predicting in-sample 3-month 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.

Table 9 provides R^2's for the out-of-sample period from April 1988 to December 1994.[18] The relative ranking of the three models that are estimated without using options is the same as for the in-sample period described in

[17] For different maturities, Cochrane and Piazzesi (2005) run regressions of yields variations on a linear combination of forward rates. For each n-year zero coupon swap rate $(n = 2, 3, 4, 5)$, they regress

$$r^{e,n}_{t,\Delta t} - r^{\Delta t}_t = \beta^n_0 + \beta^n_1 r^1_t + \beta^n_2 f^2_t + \beta^n_3 f^3_t + \beta^n_4 f^4_t + \beta^n_5 f^5_t ,$$

where $f^\tau_t := \tau r^\tau_t - (\tau - 1) r^{\tau-1}_t$ is the 1-year forward rate at time t between $t+\tau-1$ and $t+\tau$. CP_5 are the regressions described above, while CP_{10} are correspondent regressions using one period forward rates for loans between maturities that go up to 10 years. Finally, $\mathrm{CP}_{5,10}$ use only 5 one year forward rates (which begin in 0,2,4,6, and 8 years) as regressors.

[18] Recall that the models were estimated with historical data from January 1995 to February 2006, which corresponded to the availability of cap data in Datastream.

Table 8. As is the case for the in-sample analysis, when options are included in estimation, the $A_1(3)^o$ and $A_2(3)^o$ are better able to predict out-of-sample 3-month excess returns relative to their $A_1(3)$ and $A_2(3)$ counterparts that were estimated without including options. On average, the R^2s for the $A_1(3)^o$ model are larger than the R^2s for the $A_1(3)$ model by a factor of 1.3. The improvement is even larger for the $A_2(3)$ model where the R^2s are larger than the those for the $A_2(3)$ model by an average factor of 2.2. For this out-of-sample period, the $A_2(3)^o$ model is slightly better at predicting excess returns for zero coupon swaps with maturities up to 5 years, while the $A_1(3)^o$ model is slightly better at predicting excess returns for zero coupon swaps with maturities beyond 5 years. Out-of-sample, the $A_1(3)^o$ and $A_2(3)^o$ modles better predict excess returns for zero coupon swaps with maturities beyond 5 years, but the Cochrane and Piazzesi regressions CP_5 and $\text{CP}_{5,10}$ better predict excess returns for swaps with maturities of 5 years or less.

Tables 10 and 11 provide the in-sample and out-of-sample predictability for 3 month excess returns. The results are qualitatively the same as those for 1 year excess returns.

6 Linear Projection of Yields

Dai and Singleton (2002) present two challenges for dynamic term structure models that are related to predicting excess returns. This section examines whether the models we estimate satisfy these challenges.

The first challenge, which Dai and Singleton (2002) refer to as LPY(I), is to match the pattern of violations of the exectations hypothesis as in Fama and Bliss (1987) and Campbell and Shiller (1991). These papers perform the following regression

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

and find that regression coefficients $\hat{\Phi}_{1,n}$ are increasingly negative for larger maturities n.

Figure 5 provides the LPY(I) regression results for the models that we estimate.[19] However, in contrast to Dai and Singleton (2002), all of the models are generally consistent with the observed slope coefficients. For all

[19]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

of the models, the observed slope coefficients lie within a simulated 95% confidence interval. Figure 6 shows the 95% simulated confidence interval for the $A_1(3)^o$ model.

Dai and Singleton (2002) refer to the second challenge that they pose as LPY(II). This challenges states that the projection of risk-adjusted changes in swap rates onto the slope of the yield curve should recover a regression coefficient of 1. That is, if the risk premium in the model is correct, then the risk premium adjusted regression

$$\underbrace{r_{t+\Delta t}^{n-\Delta t} - r_t^n + \frac{\Delta t}{n-\Delta t}\mathbb{E}_t\left[r_{t,\Delta t}^{e,n}\right]}_{\text{PACY}_{t,\Delta t}^n} = \Phi_{0,n} + \Phi_{1,n}\underbrace{\frac{\Delta t}{n-\Delta t}\left(r_t^n - r_t^\Delta\right)}_{\text{SLOPE}_t^n} + \varepsilon_t^n\,, \tag{5}$$

should produce a regression coefficient $\hat{\Phi}_{1,n} = 1$.

We find that the combination of small sample size and near unit roots in swap rates results in a small, but non-negligible, bias in the regression coefficients. In order to better understand the source of this bias, note that the regression coefficient in equation (5) is $\hat{\Phi}_{1,n} = U/V$, where

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

and $\overline{\text{PACY}^n}$ and $\overline{\text{SLOPE}^n}$ are the sample averages. In general, $\mathbb{E}[U] \neq \mathbb{E}[V]$. 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

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

minimize the effect from bootstrapping the zero coupon yield curve. The results for 1 month changes are qualitatively similar.

In general, for $k < m$, the residual $\varepsilon_{k\Delta t} = \mathbb{E}_{k\Delta t}\left[\mathrm{PACY}^n_{k\Delta t,\Delta t}\right] - \mathrm{SLOPE}^n_{k\Delta t}$ may be correlated with $\mathrm{SLOPE}^n_{k\Delta t}$ (and is likely to be more correlated when $\mathrm{SLOPE}^n_{k\Delta t}$ is more nearly stationary). Therefore $\mathbb{E}\left[\mathrm{PACY}^n_{k\Delta t,\Delta t}\mathrm{SLOPE}^n_{m\Delta t}\right] \neq \mathbb{E}\left[\mathrm{SLOPE}^n_{k\Delta t}\mathrm{SLOPE}^n_{m\Delta t}\right]$.[20]

The bias can be approximated by a second order Taylor series expansion:

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

Though cumbersome, this 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 7 shows the model LPY(II) regression results adjusted for the small sample bias. The $A_1(3)^o$ and $A_2(3)^o$ models estimated with options dominate the $A_0(3)$, $A_1(3)$, and $A_2(3)$ models estimated without options. Although we do not recover an exact regression coefficient of one, this value is quite near the center of the simulated confidence intervals for the $A_0(3), A_1()^o$, and $A_2(3)^o$ models. Figure 8 shows the simulated 95% confidence interval for the $A_1(3)^o$ model. The stochastic volatility models without options nearly follow the 95% lower confidence bound.

7 Conclusion

Theory suggests that 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 risk factors that drive 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 estimated using interest rate option prices are better able to predict excess returns for long-term swaps over short-term swaps, both in- and out-of-sample. We also find that, with a correction for small sample bias, the models with stochastic volatility successfully capture the failure of the expectations hypothesis and match regressions of returns on the slope of the yield curve.

The question remains, what elements of the risk premium do including options help to better identify? In dynamic term structure models with

[20] 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.

stochastic volatility, interest rate cap prices depend on the price of volatility risk. Therefore, the price of volatility risk is one element of the risk premium that may depend on whether options are used in estimation.

As a measure of the price of volatility risk we use the difference between the risk neutral expected zero coupon bond volatility and the expected zero coupon bond volatility provides a measure of the price of volatility risk. Figures 9 and 10 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.[21] When options are not included in estimation, this measure of the price of volatility risk is more negative in the $A_1(3)$ model, and large and positive for the $A_2(3)$ model. This measure of the price of volatility risk is essentially zero for the $A_2(3)^o$ model that is estimated with options and small but negative for the $A_1(3)^o$ model that is estimated with options. Therefore, including options affects the estimated price of volatility risk. 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.

References

Ruslan Bikbov and Mikhail Chernov. Term structure and volatility: Lessons from the eurodollar markets. Working Paper, 7 June 2004.

Fischer Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2):167–179, January-March 1976.

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.

[21] The long run expected volatility of the τ-year zero coupon bond yield 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 bond yield is

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

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.

Patrick Cheridito, Damir Filipovic, and Robert L. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, September 2006.

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

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.

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.

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

Bin Han. Stochastic volatilities and correlations of bond yields. Working Paper, October 2004.

Scott Joslin. Completeness in fixed income markets. Working Paper, Stanford University, 2005.

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.

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

A Detailed Model Specifications

We estimate 3-factor affine term structure models. The state vector X follows an diffusion process with affine drift and variance,

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

We have used the notation $\Delta[\cdot]$ to denote a square matrix with its vector argument along the diagonal. The pricing kernel M also follows a diffusion process

$$dM_t = -M_t\, r_t\, dt - M_t\, \Lambda_t^\top d W_t\,.$$

The drift, r, of the pricing kernel is an affine function of the state vector

$$r_t := \rho_0 + \rho_1 \cdot X_t\,,$$

The volatility, Λ, of the pricing kernel has an extended affine form

$$\Lambda_t := \left(\sqrt{\Delta\left[\alpha + \beta X_t\right]}\right)^{-1} \left[\left(\mathcal{K}_0^{\mathbb{P}} - \mathcal{K}_0\right) + \left(\mathcal{K}_1^{\mathbb{P}} - \mathcal{K}_1\right) X_t\right]\,.$$[22, 23]

We estimate $A_M(3)$ models,[24] where $M = 0, 1, 2$ is the number of factors in the state vector X that have stochastic volatility. For example, in the $A_0(3)$ model, $\beta = 0$ so that none of the factors have stochastic volatility. For each model, Dai and Singleton (2000) and Cheridito et al. (2006) identify the necessary restrictions required to ensure that the stochastic processes are admissable, the parameters are identified, and the physical and risk neutral measures are equivalent. The full specifications of the $A_0\,(3)$, $A_1\,(3)$, and $A_2\,(3)$ are described below.

---

[23] We use an extended affine market price of risk introduced by Cheridito et al. (2006) as a generalization of the essentially affine market price of risk used in Duffee (2002).

[23] The dynamics of the state vector X under the martingale pricing measure $\mathbb{Q}$ defined by

$$\left.\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\right|_t := e^{\int_0^t r_u\, du}\, \frac{M_t}{M_0} = e^{-\frac{1}{2}\int_0^t \Lambda_u^\top \Lambda_u\, du - \int_0^t \Lambda_u^\top d\, W_u}\,,$$

are

$$dX_t \;=\; \left[\mathcal{K}_0 + \mathcal{K}_1 X_t\right] dt + \sqrt{\Delta\left[\alpha + \beta X_t\right]}\, d W_t^{\mathbb{Q}}\,,$$

where

$$W_t^{\mathbb{Q}} := W_t + \int_0^t \Lambda_u\, du\,,$$

is a Brownian motion under $\mathbb{Q}$.

[24] The models were introduced in Dai and Singleton (2000).

$A_0(3)$ Model Specification

$$\begin{aligned}
\mathcal{K}_1^{\mathbb{P}} &= \begin{bmatrix} \mathcal{K}_{1,11}^{\mathbb{P}} & 0 & 0 \\ \mathcal{K}_{1,21}^{\mathbb{P}} & \mathcal{K}_{1,22}^{\mathbb{P}} & 0 \\ \mathcal{K}_{1,31}^{\mathbb{P}} & \mathcal{K}_{1,32}^{\mathbb{P}} & \mathcal{K}_{1,33}^{\mathbb{P}} \end{bmatrix}, \quad \mathcal{K}_0^{\mathbb{P}} = \begin{bmatrix} \mathcal{K}_{0,1}^{\mathbb{P}} \\ \mathcal{K}_{0,2}^{\mathbb{P}} \\ \mathcal{K}_{0,3}^{\mathbb{P}} \end{bmatrix}, \\
\mathcal{K}_1 &= \begin{bmatrix} \mathcal{K}_{1,11} & 0 & 0 \\ \mathcal{K}_{1,21} & \mathcal{K}_{1,22} & 0 \\ \mathcal{K}_{1,31} & \mathcal{K}_{1,32} & \mathcal{K}_{1,33} \end{bmatrix}, \quad \mathcal{K}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \\
\beta &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},
\end{aligned}$$

$$\rho_{1,1} \geq 0\,, \quad \rho_{1,2} \geq 0\,, \quad \rho_{1,3} \geq 0\,.$$

$A_1(3)$ Model Specification

$$\begin{aligned}
\mathcal{K}_1^{\mathbb{P}} &= \begin{bmatrix} \mathcal{K}_{1,11}^{\mathbb{P}} & 0 & 0 \\ \mathcal{K}_{1,21}^{\mathbb{P}} & \mathcal{K}_{1,22}^{\mathbb{P}} & \mathcal{K}_{1,23}^{\mathbb{P}} \\ \mathcal{K}_{1,31}^{\mathbb{P}} & \mathcal{K}_{1,32}^{\mathbb{P}} & \mathcal{K}_{1,33}^{\mathbb{P}} \end{bmatrix}, \quad \mathcal{K}_0^{\mathbb{P}} = \begin{bmatrix} \mathcal{K}_{0,1}^{\mathbb{P}} \\ \mathcal{K}_{0,2}^{\mathbb{P}} \\ \mathcal{K}_{0,3}^{\mathbb{P}} \end{bmatrix}, \\
\mathcal{K}_1 &= \begin{bmatrix} \mathcal{K}_{1,11} & 0 & 0 \\ \mathcal{K}_{1,21} & \mathcal{K}_{1,22} & \mathcal{K}_{1,23} \\ \mathcal{K}_{1,31} & \mathcal{K}_{1,32} & \mathcal{K}_{1,33} \end{bmatrix}, \quad \mathcal{K}_0 = \begin{bmatrix} \mathcal{K}_{0,1} \\ 0 \\ 0 \end{bmatrix}, \\
\beta &= \begin{bmatrix} 1 & 0 & 0 \\ \beta_{2,1} & 0 & 0 \\ \beta_{3,1} & 0 & 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
&\mathcal{K}_{0,1}^{\mathbb{P}} \geq \frac{1}{2}\,, \quad \mathcal{K}_{0,1} \geq \frac{1}{2}\,, \\
&\beta_{2,1} \geq 0\,, \quad \beta_{3,1} \geq 0\,, \\
&\rho_{1,2} \geq 0\,, \quad \rho_{1,3} \geq 0\,.
\end{aligned}$$

$A_2(3)$ Model Specification

$$\begin{aligned}
\mathcal{K}_1^{\mathbb{P}} &= \begin{bmatrix} \mathcal{K}_{1,11}^{\mathbb{P}} & \mathcal{K}_{1,12}^{\mathbb{P}} & 0 \\ \mathcal{K}_{1,21}^{\mathbb{P}} & \mathcal{K}_{1,22}^{\mathbb{P}} & 0 \\ \mathcal{K}_{1,31}^{\mathbb{P}} & \mathcal{K}_{1,32}^{\mathbb{P}} & \mathcal{K}_{1,33}^{\mathbb{P}} \end{bmatrix}, \quad \mathcal{K}_0^{\mathbb{P}} = \begin{bmatrix} \mathcal{K}_{0,1}^{\mathbb{P}} \\ \mathcal{K}_{0,2}^{\mathbb{P}} \\ \mathcal{K}_{0,3}^{\mathbb{P}} \end{bmatrix}, \\
\mathcal{K}_1 &= \begin{bmatrix} \mathcal{K}_{1,11} & \mathcal{K}_{1,12} & 0 \\ \mathcal{K}_{1,21} & \mathcal{K}_{1,22} & 0 \\ \mathcal{K}_{1,31} & \mathcal{K}_{1,32} & \mathcal{K}_{1,33} \end{bmatrix}, \quad \mathcal{K}_0 = \begin{bmatrix} \mathcal{K}_{0,1} \\ \mathcal{K}_{0,2} \\ 0 \end{bmatrix}, \\
\beta &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \beta_{3,1} & \beta_{3,2} & 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
&\mathcal{K}_{0,1}^{\mathbb{P}} \geq \frac{1}{2}, \quad \mathcal{K}_{0,1} \geq \frac{1}{2}, \quad \mathcal{K}_{0,2}^{\mathbb{P}} \geq \frac{1}{2}, \quad \mathcal{K}_{0,2} \geq \frac{1}{2}, \\
&\beta_{3,1} \geq 0, \quad \beta_{3,2} \geq 0, \\
&\rho_{1,3} \geq 0.
\end{aligned}$$

B Detailed Estimation Procedure

We estimate all the models using quasi-maximum likelihood in a procedure similar to Duffee (2002) and Dai and Singleton (2002). 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:

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

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

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

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

$$\dot{f} = \left[\begin{array}{cc} K_1 & 0 \\ \Delta & I_N \otimes K_1 + K_1 \otimes I_N \end{array} \right] f + \left[\begin{array}{c} K_0 \\ \mathrm{vec}(H_0) \end{array} \right]$$

Where Δ is an $(N^2 \times N)$ matrix with $\Delta_{i,j} = \mathrm{vec}(H_{1,\cdot,\cdot,i})_j$. 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.

C Tables and Figures

	2 Year	3 Year	4 Year	5 Year
Slope Only	6.21	8.23	8.96	9.00
Slope and Cap Implied Volatility	17.29	27.52	34.26	38.06
All Yields	32.61	41.50	48.03	52.62

Table 1: Regression of Excess Returns.
This table shows the R^2 from regressions of (overlapping) one year excess returns for 2- to 5-year zero coupon swaps. The sample period is January 1995 to February 2006. In the M-year regression, slope was taken as the difference between the M-year zero coupon swap rate and 1-year zero coupon swap rate. The M-year cap implied volatility was used in the M-year regression.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
$K^{1P}_{1,1}$	-1.4 (0.486)	-0.21 (0.339)	-0.307 (0.319)	-0.244 (0.371)	-2.35 (0.649)
$K^{1P}_{1,2}$	0.448 (0.459)	0	0	1.06 (1.16)	2.94 (0.857)
$K^{1P}_{1,3}$	-0.108 (0.108)	0	0	0	0
$K^{1P}_{2,1}$	0.781 (0.497)	-1.96 (0.501)	-2.23 (0.45)	0.193 (0.157)	0.69 (0.458)
$K^{1P}_{2,2}$	-0.893 (0.47)	-0.962 (0.429)	-0.441 (0.3)	-1.29 (0.477)	-1.06 (0.577)
$K^{1P}_{2,3}$	0.259 (0.124)	-0.58 (1.39)	-0.801 (2)	0	0
$K^{1P}_{3,1}$	-3.27 (0.728)	-1.88 (4.98)	-4 (9.69)	0.817 (1.8)	2.11 (0.745)
$K^{1P}_{3,2}$	1.06 (0.652)	-0.438 (1.24)	-0.341 (1)	-3.42 (6.83)	-3.61 (0.891)
$K^{1P}_{3,3}$	-0.46 (0.172)	-0.767 (0.51)	-2.19 (0.813)	-0.511 (0.164)	-0.321 (0.143)
K^{0P}_{1}	1.04 (1.98)	1.38 (1.9)	2.28 (2.53)	1.97 (21.7)	1.61 (5.11)
K^{0P}_{2}	-1.31 (1.99)	-0.0177 (3.72)	7.38 (4.99)	0.614 (10.2)	0.5 (3.2)
K^{0P}_{3}	-0.184 (2.95)	-1.71 (9.05)	-1.8 (12.2)	0	0
$K^{1Q}_{1,1}$	-1.3 (0.043)	-0.62 (0.0128)	-0.568 (0.0111)	-0.0864 (0.0634)	-1.29 (0.0669)
$K^{1Q}_{1,2}$	0	0	0	0.484 (0.377)	1.64 (0.245)
$K^{1Q}_{1,3}$	0	0	0	0	0
$K^{1Q}_{2,1}$	-0.0947 (0.0443)	-2.06 (0.334)	-2.11 (0.336)	0.176 (0.109)	0.0652 (0.0294)
$K^{1Q}_{2,2}$	-0.0298 (0.0013)	-0.939 (0.42)	-0.236 (0.354)	-1.35 (0.0748)	-0.756 (0.0601)
$K^{1Q}_{2,3}$	0	-0.723 (1.7)	-1.21 (2.98)	0	0
$K^{1Q}_{3,1}$	-4 (0.159)	-1.61 (4.27)	-2.88 (7)	1.1 (2.31)	2.95 (0.817)
$K^{1Q}_{3,2}$	1.38 (0.0544)	-0.592 (1.52)	-0.168 (0.492)	-4 (8.12)	-3.64 (0.811)
$K^{1Q}_{3,3}$	-0.681 (0.014)	-0.502 (0.413)	-1.44 (0.355)	-0.653 (0.0117)	-0.0819 (0.00106)
K^{0Q}_{1}	0	2.81 (0.682)	3.62 (0.769)	1.67 (3.11)	0.5 (1.76)
K^{0Q}_{2}	0	0	0	1.29 (5.39)	3.42 (1.02)
K^{0Q}_{3}	0	0	0	-1.37 (13.9)	-1.23 (2.84)
$\beta_{1,1}$	0	1	1	1	1
$\beta_{1,2}$	0	0.0638 (0.0361)	0.149 (0.0698)	0	0
$\beta_{1,1}$	0	4.49 (21.3)	5.84 (28.6)	0 (0.0612)	0 (0.0478)
$\beta_{1,1}$	0	0	0	1	1
$\beta_{1,1}$	0	0	0	0.74 (3.12)	0.411 (0.253)
ρ^0	0.111 (0.00229)	0.0478 (0.0186)	0.0166 (0.0111)	-0.0148 (0.0871)	0.0684 (0.0422)
ρ^1_1	0.00241 (2.28e-4)	0.000734 (1.48e-4)	0.000811 (1.24e-4)	-8.28e-5 (6.66e-5)	-0.000995 (9.91e-5)
ρ^1_2	0.00022 (2.62e-4)	0.00454 (3.22e-4)	0.00309 (3.88e-4)	0.000897 (2.25e-4)	0.00113 (1.59e-4)
ρ^1_3	0.00559 (8.89e-5)	0.000252 (5.55e-4)	0.000487 (0.00121)	0.00191 (0.00395)	0.00236 (6.53e-4)
M	0	1	1	2	2
N	3	3	3	3	3
mean LLK	41.6	41.7	74.6	41.6	74.7

Table 2: Parameter Estimates.
This table presents all parameter values for the different affine term structure models we estimate. Standard errors are 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 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. If a parameter is reported as 0 or 1, it is restricted to be so by the identification and existence conditions in Dai and Singleton (2000) and Cheridito et al. (2006).

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
3 Month	0.0	0.0	0.0	0.0	0.0
6 Month	7.1	7.1	6.8	7.1	6.8
1 Year	9.9	9.9	9.3	10.0	9.3
2 Year	0.0	0.0	0.0	0.0	0.0
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
10 Year	0.0	0.0	0.0	0.0	0.0

Table 3: Pricing Errors in BPS for 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
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 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.

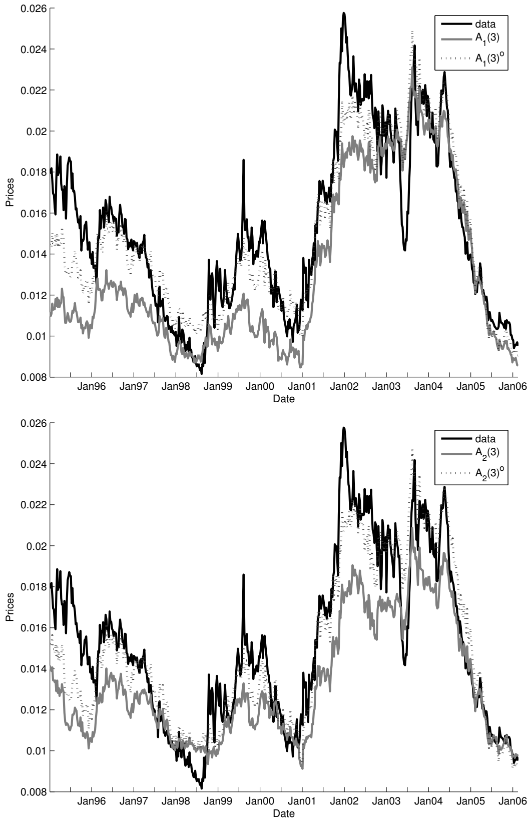

Figure 1: Cap Prices

The top figure plots the 3-year at-the-money cap prices in the $A_1(3)$ model (solid grey line) and the $A_1(3)^o$ model (dashed line). The actual prices are plotted with a solid black line. The bottom figure plots the 3-year at-the-money cap prices in the $A_2(3)$ model (solid grey line) and the $A_2(3)^o$ model (dashed line). The actual prices are plotted with a solid black line. 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.

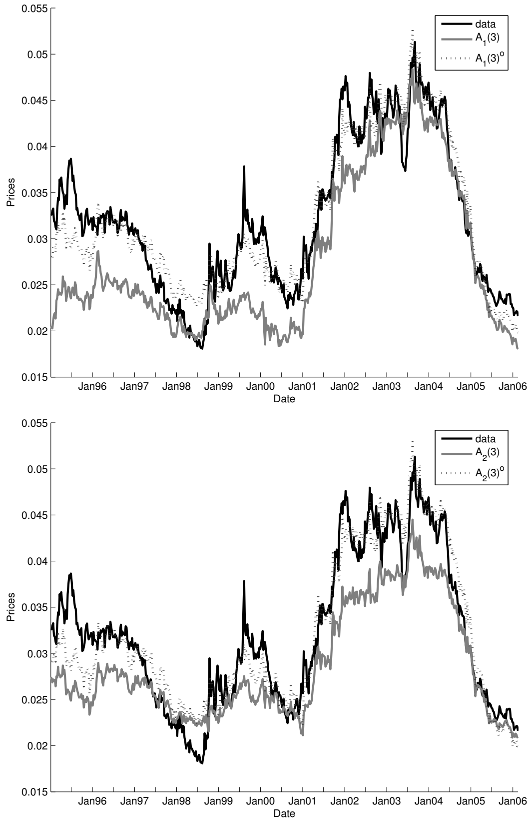

Figure 2: Cap Prices

The top figure plots the 5-year at-the-money cap prices in the $A_1(3)$ model (solid grey line) and the $A_1(3)^o$ model (dashed line). The actual prices are plotted with a solid black line. The bottom figure plots the 5-year at-the-money cap prices in the $A_2(3)$ model (solid grey line) and the $A_2(3)^o$ model (dashed line). The actual prices are plotted with a solid black line. 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.

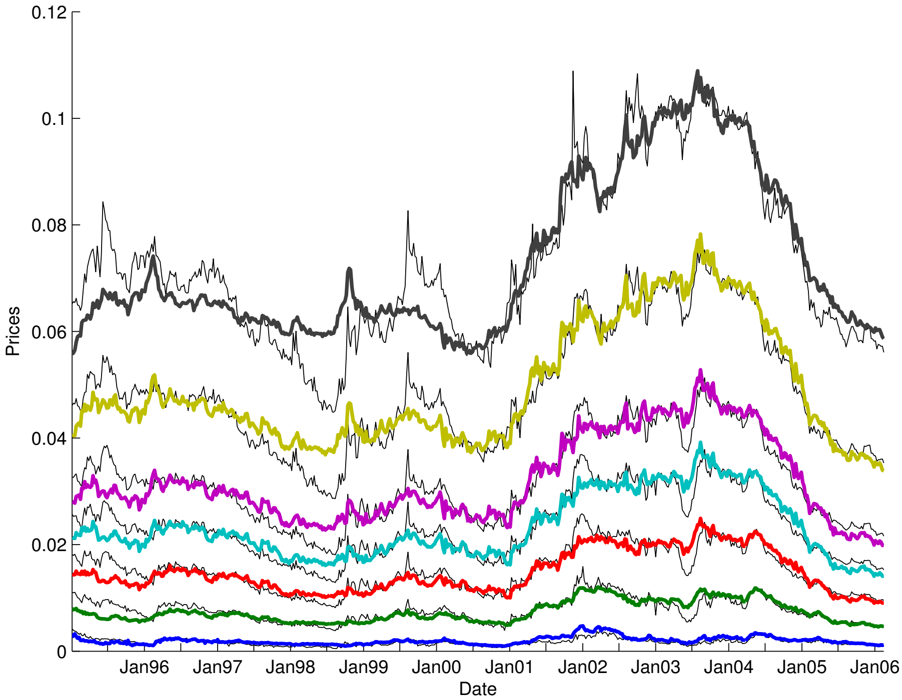

Figure 3: Cap Prices for $A_1(3)^o$ model

This figure plots cap prices for 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps. The think lines indicate data 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.

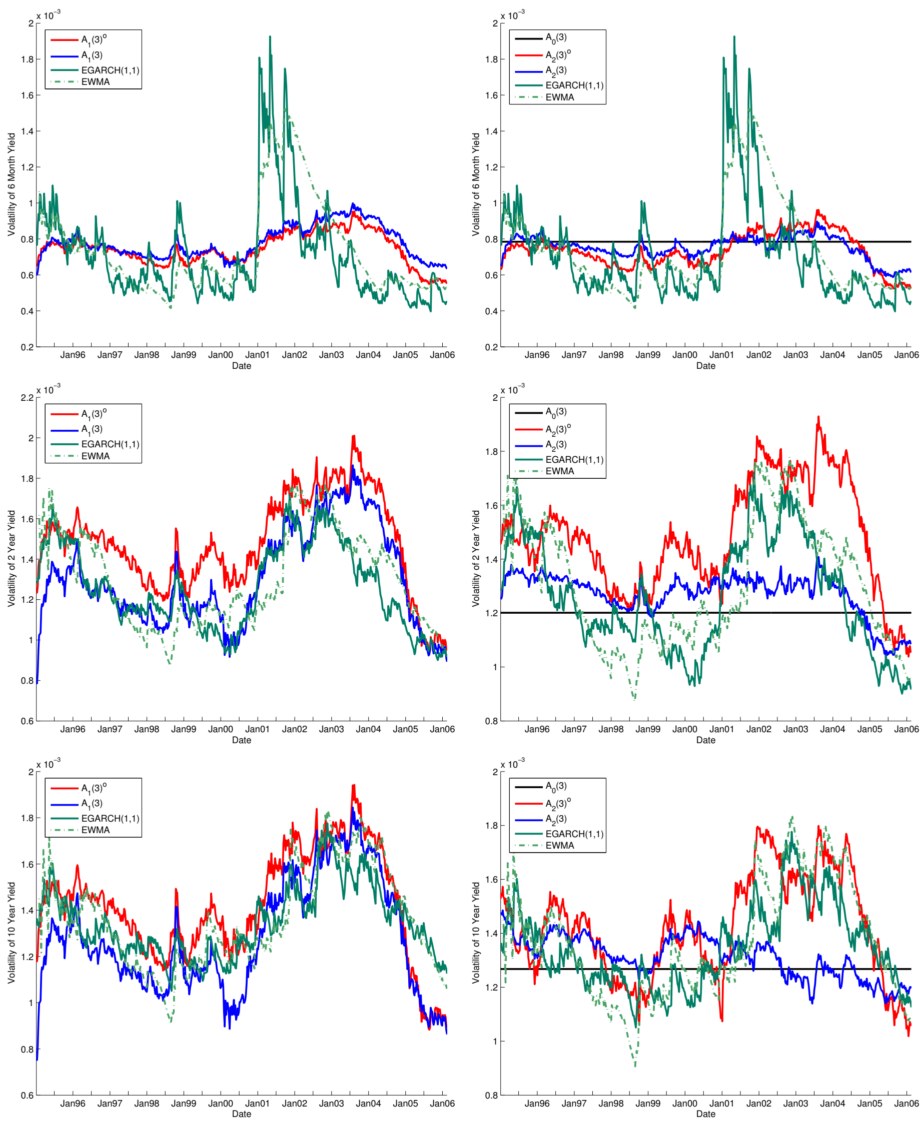

Figure 4: Realized Volatility

These figures plot weekly model conditional volatility of (from top to bottom) 6-month, 2-year, and 10-year zero coupon swap rates 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 plots on the left show the conditional volatility in the $A_1(3)$ and $A_1(3)^o$ models. The plots on the right show the conditional volatility in the $A_0(3)$, $A_1(3)$, and $A_1(3)^o$ models.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
6 Month	0.0	19.2	28.9	39.1	30.1
1 Year	0.0	50.8	56.3	58.3	52.9
2 Year	0.0	75.0	77.0	63.2	66.9
3 Year	0.0	83.0	81.5	39.0	70.6
4 Year	0.0	84.4	81.4	15.4	71.9
5 Year	0.0	84.1	79.3	-2.6	69.3
7 Year	0.0	84.3	77.4	-21.2	66.4
10 Year	0.0	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 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	0.0	33.2	41.8	35.1	40.3
1 Year	0.0	55.8	62.7	54.5	63.4
2 Year	0.0	74.0	76.0	50.1	76.1
3 Year	0.0	76.7	75.6	29.6	75.6
4 Year	0.0	77.2	74.1	9.9	74.0
5 Year	0.0	77.4	73.0	-3.4	72.0
7 Year	0.0	80.6	74.2	-18.7	71.4
10 Year	0.0	80.7	75.8	-18.9	71.7

Table 6: 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	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 7: 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	CP_5	CP_{10}	$CP_{5,10}$
2 Yr	-15.2	10.8	2.2	12.1	5.8	36.4	45.3	36.1
3 Yr	-1.1	16.5	15.4	16.5	21.9	42.5	50.4	40.8
4 Yr	8.6	20.0	23.4	17.0	30.1	47.8	54.6	44.6
5 Yr	15.3	21.8	28.5	16.8	34.7	51.5	57.5	46.8
6 Yr	20.2	22.5	32.0	16.4	37.3		59.0	47.0
7 Yr	24.5	21.9	34.3	15.7	38.6		60.0	46.4
8 Yr	26.9	21.6	35.6	15.4	39.1		60.0	45.0
9 Yr	29.3	21.0	36.6	15.0	39.4		59.9	43.5
10 Yr	31.1	20.2	37.1	14.6	39.1		59.8	42.0

Table 8: In-Sample 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 in-sample model implied returns. 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$	CP_5	CP_{10}	$CP_{5,10}$
2 Yr	-11.7	12.3	1.0	10.1	1.3	23.4	-75.0	30.1
3 Yr	-3.3	18.9	12.0	13.6	13.9	29.1	-56.1	26.2
4 Yr	4.4	20.9	18.0	12.9	19.5	33.2	-40.5	22.5
5 Yr	10.9	21.7	23.0	12.1	23.6	36.4	-26.9	23.2
6 Yr	16.2	19.3	25.4	10.4	24.5		-19.3	18.5
7 Yr	21.3	18.2	28.7	9.5	26.3		-14.6	15.0
8 Yr	24.5	18.3	31.2	9.4	28.1		-13.4	12.2
9 Yr	27.2	16.7	32.4	8.6	28.3		-12.1	8.6
10 Yr	29.3	15.1	33.1	8.0	28.1		-10.8	5.6

Table 9: Out-of-Sample 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 out-of-sample model implied returns. 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
2 Yr	4.2	6.7	9.2	5.7	9.9
3 Yr	8.7	7.7	12.4	6.6	14.3
4 Yr	11.6	8.3	14.4	6.6	16.3
5 Yr	13.4	8.5	15.3	6.4	17.1
6 Yr	14.4	8.6	15.8	6.3	17.4
7 Yr	15.2	8.4	15.9	6.0	17.3
8 Yr	15.6	8.4	16.0	5.9	17.3
9 Yr	15.9	8.3	15.9	5.7	17.0
10 Yr	16.1	8.1	15.7	5.5	16.7

Table 10: In-Sample 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 in-sample 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.

	$A_0(3)$	$A_1(3)$	$A_1(3)^o$	$A_2(3)$	$A_2(3)^o$
2 Yr	1.5	4.8	5.1	2.8	4.4
3 Yr	3.1	5.8	6.3	3.3	5.6
4 Yr	4.1	6.1	6.7	3.1	5.6
5 Yr	5.1	6.1	7.1	2.9	5.8
6 Yr	5.2	5.8	6.7	2.5	4.9
7 Yr	5.3	5.5	6.5	2.2	4.3
8 Yr	5.3	5.6	6.4	2.0	4.0
9 Yr	5.1	5.3	6.0	1.7	3.4
10 Yr	4.8	5.0	5.7	1.5	3.0

Table 11: Out-of-Sample 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 out-of-sample 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.

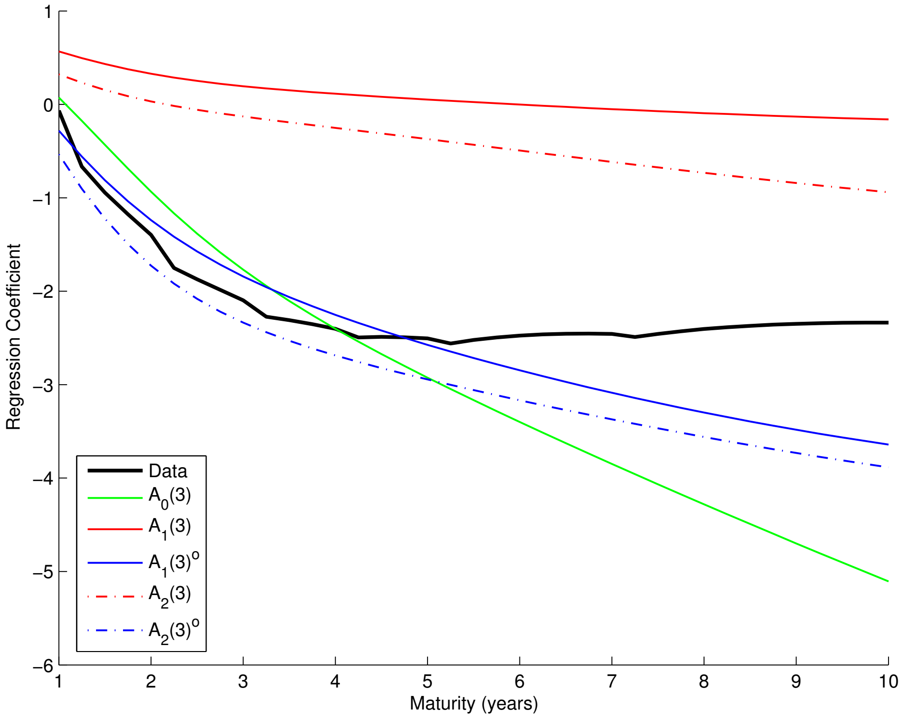

Figure 5: Regression Coefficients from Linear Projection on Yields This figure shows the regression coefficients of the Campbell-Shiller regression $R_{t+1}^{n-1} - R_t^n$ on slope, $(R_t^n - r_t)/(n-1)$. The model values are simulated mean regression coefficients. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ 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.

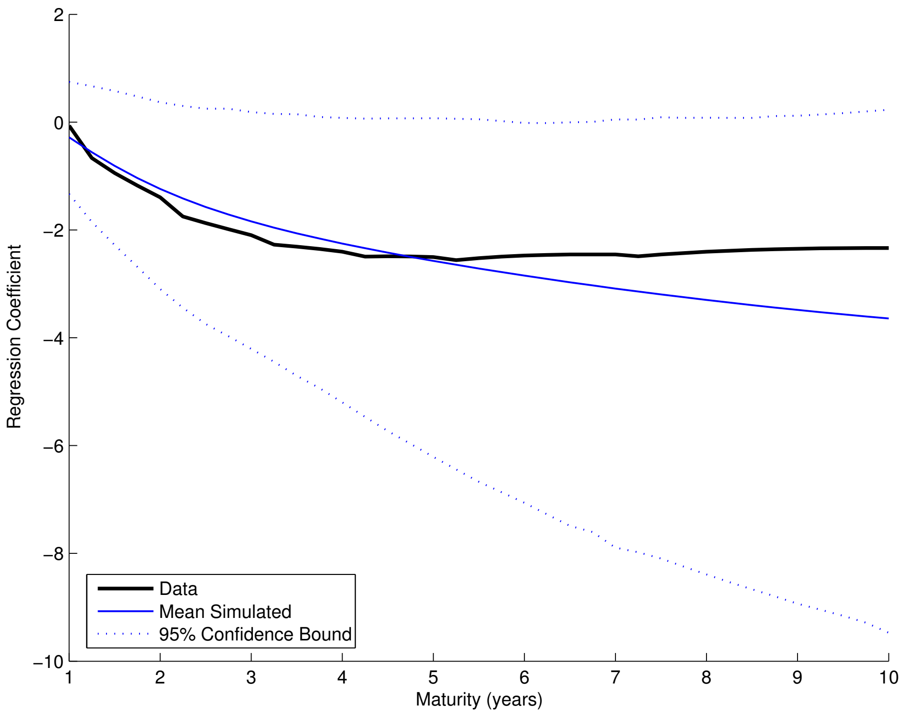

Figure 6: Confidence Interval for Linear Projection on Yields
This figure shows the sample regression coefficient of the Campbell-Shiller regression $R_{t+1}^{n-1} - R_t^n$ on slope, $(R_t^n - r_t)/(n-1)$ for the $A_1(3)^o$ model. The dotted line provides the confidence interval computed from simulation. 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.

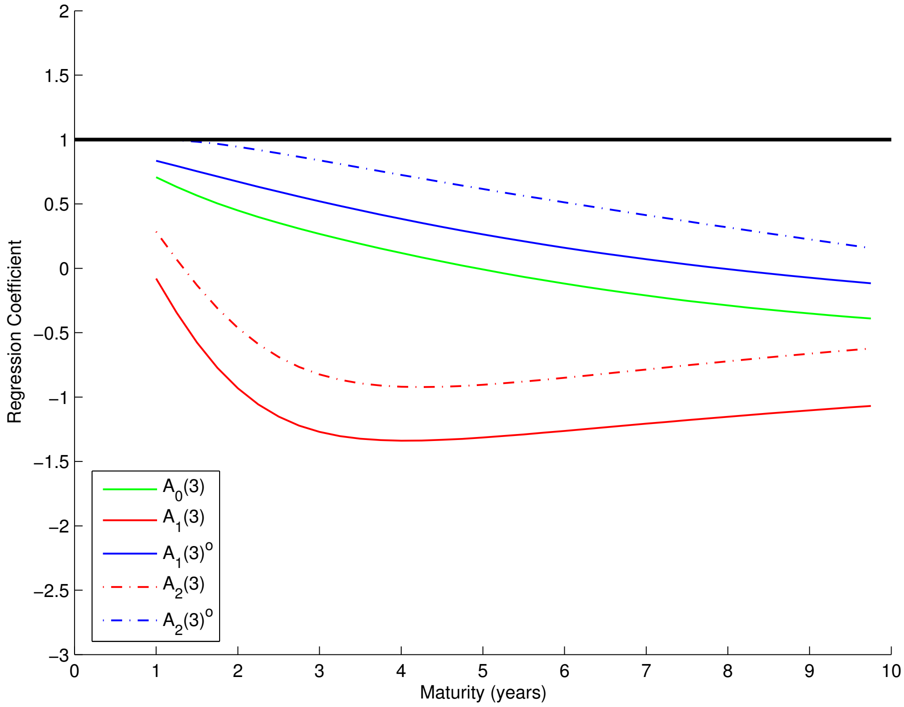

Figure 7: Risk Premium Adjusted Linear Projection on Yields
This figure shows the regression coefficient from the projection risk premium adjusted excess returns on the slope, $(R^n_t - r_t)/(n-1)$. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-month, 2-year, and 10-year swap 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.

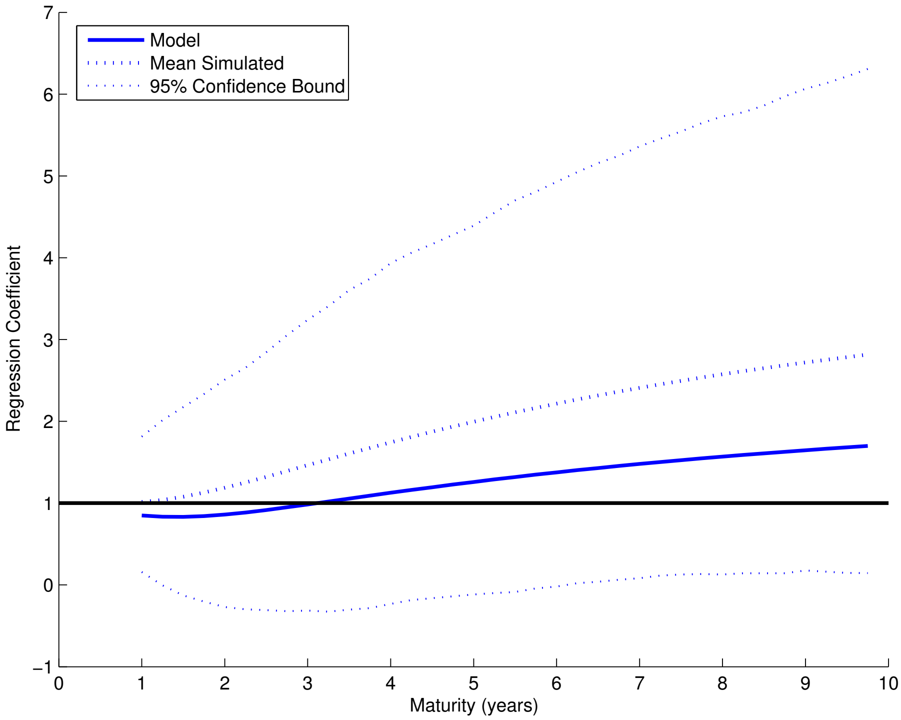

Figure 8: Confidence Interval for LPY(II)

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 $A_1(3)^o$ 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.

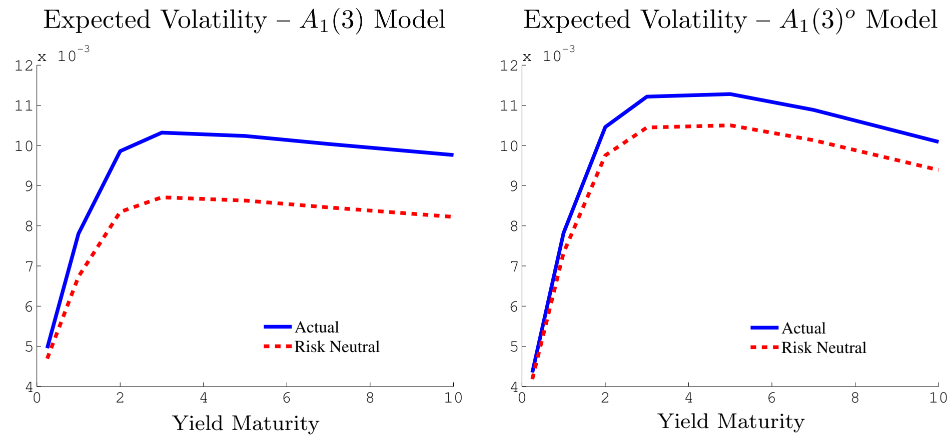

Figure 9: Actual and Risk Neutral Expected Volatility
These figures show the long run expected volatility of zero coupon yields with maturities up to 10 years. The plot on the left shows the expected long run volatility in the $A_1(3)$ model that is estimated without using options. The plot on the right 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.

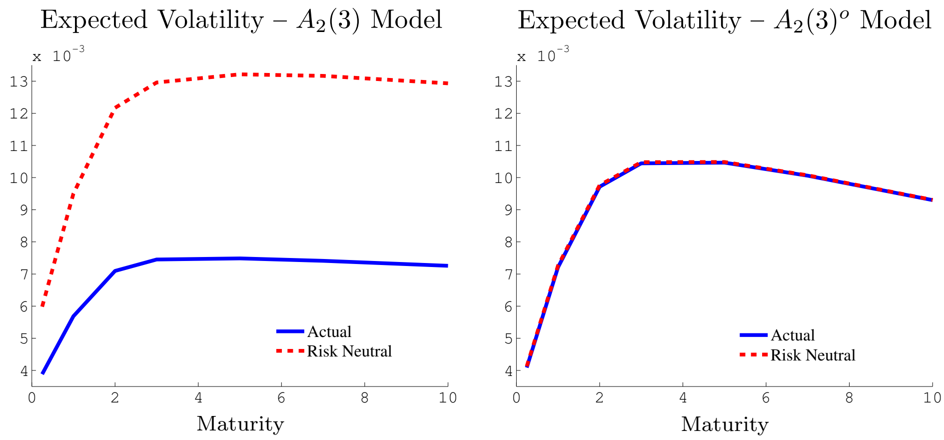

Figure 10: Actual and Risk Neutral Expected Volatility
These figures show the long run expected volatility of zero coupon yields with maturities up to 10 years. The plot on the left shows the expected long run volatility in the $A_2(3)$ model that is estimated without using options. The plot on the right 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.