

**Economics Research Group
IBMEC Business School – Rio de Janeiro
<http://professores.ibmecrj.br/erg>**

IBMEC RJ ECONOMICS DISCUSSION PAPER 2005-04
Full series available at <http://professores.ibmecrj.br/erg/dp/dp.htm>.

DO OPTIONS CONTAIN INFORMATION ABOUT EXCESS BOND RETURNS?

**Caio Almeida
Jeremy J. Graveline
Scott Joslin**

Do Options Contain Information About Excess Bond Returns?*

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

First Draft: June 3, 2005

Current Draft: November 18, 2005

COMMENTS WELCOME

PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

There is strong empirical evidence that risk premia in long-term interest rates are time-varying. These risk premia critically depend on interest rate volatility, yet existing research has not examined the impact of time-varying volatility on excess returns for long-term bonds. To address this issue, we incorporate interest rate option prices, which are very sensitive to interest rate volatility, into a dynamic model for the term structure of interest rates. We estimate three-factor affine term structure models using both swap rates and interest rate cap prices. When we incorporate option prices, the model better captures interest rate volatility and is better able to predict excess returns for long-term swaps over short-term swaps, both in- and out-of-sample. Our results indicate that interest rate options contain valuable information about risk premia and interest rate dynamics that cannot be extracted from interest rates alone.

*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

In a regression framework, Fama and Bliss (1987) demonstrate that expected excess bond returns are both predictable and time-varying. Campbell and Shiller (1991) present further evidence from regressions that risk premiums on long-term bonds are also time varying. Recently, Duffee (2002) and Dai and Singleton (2002) have shown that dynamic term structure models with a flexible specification of the market price of interest rate risk can capture this variation in expected returns. However, the expected excess returns for long-term bonds depends on both the price of interest rate risk as well as the amount of interest rate volatility, yet comparatively little research attention has been focused on the impact of time-varying volatility on expected excess returns.

With few exceptions, previous research has not included interest rate options when estimating dynamic term structure models and therefore has not exploited the additional information about interest rate volatility that may be contained in these option prices. In this paper we estimate arbitrage-free dynamic term structure models jointly on both swap rates and the prices of interest rate caps. We use quasi-maximum likelihood to estimate three-factor affine term structure models with 0, 1, or 2 factors having stochastic volatility.¹ In order to make estimation with cap prices computationally feasible, we build on the work of Jarrow and Rudd (1982) and develop a computa-

¹See Dai and Singleton (2000) for a detailed specification of the $A_M(N)$ affine term structure models that we estimate in this paper.

tionally efficient method for computing cap prices that is well-suited to estimation. When we incorporate information in option prices, we significantly improve the model's ability to price interest options without impairing its ability to capture the term structure of interest rates. More importantly, the model's that are estimated with options are dramatically better at predicting excess returns for long-term swaps over short-term swaps, both in- and out-of-sample.

Previous papers that have used both interest rates and interest rate options in estimation have focused on accurately pricing both interest rates and options. Umantsev (2002) estimates affine models jointly on both swaps and swaptions and analyzes the volatility structure of these 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 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 affects 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 2 describes the dynamic term structure models, data, and our estimation procedure. Section 3 presents the cross-sectional fit to swap rates and cap prices. Section 4 examines the fit to swaption implied volatilities and to historical estimates

of conditional volatility. Section 5 compares the estimated models' ability to predict excess returns and Section 6 concludes. Technical details, and all tables and figures are contained in the appendix.

1 Excess Returns in Fixed Income Markets

Any bond held for a period less than its maturity will have a risky return. For example, although the 5-year interest rate is known, the return on a 5-year bond that is sold in one year is uncertain and risky. Economic reasoning suggests that investors may demand a premium for holding this risk. Interest rate volatility is one measure of the amount of such risk that a bond is exposed to, and in this section we use regression analysis to test whether interest rate volatility explains variation in bond returns.

Specifically, defining:

$$\begin{aligned}
 p_t^n &= \text{price at time } t \text{ of } n \text{ year zero coupon bond,} \\
 r_t^n &= n\text{-year continuously compounded yield} \\
 &= -\frac{1}{n} \log(p_t^n).
 \end{aligned}$$

The log excess return for holding an n -year bond for one year is then:

$$\begin{aligned}
 r_{t+1}^{e,n} &= \log(p_{t+1}^{n-1}) - \log(p_t^n) - r_t^1 \\
 &= -(n-1)r_{t+1}^{n-1} + nr_t^n - r_t^1
 \end{aligned}$$

Previous papers have shown that the current term structure of interest rates can be used to predict excess bond returns. For instance, Fama and Bliss (1987) and Campbell and Shiller (1991) provide evidence that the slope of the yield curve explains variation in excess returns. Cochrane and Piazzesi (2005) argue that the entire yield curve provides valuable information for explaining variation in risk premia.

If investors demand a premium for holding long term bonds with a risky return, then interest rate volatility may provide additional predictive power in a regression. As a measure of volatility, we use interest rate cap data. 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 our regression analysis, we use the Black implied volatility of at-the-money caps as a measure of the unobserved true volatility.² The implied volatility from at-the-money caps provides a forward looking measure of volatility that incorporates risk preferences.

As a preliminary test of this hypothesis, we regress the one year excess returns of 2- to 5-year bonds 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,

²See Section 2 for a detailed description of the data.

3. one to five year zero rates.

We report the R^2 from the regressions using 483 weekly observations from June 1995 to March 2004 in Table 1. The results indicate that indeed including the cap implied volatility increases the amount of variation which is explained from just using the slope alone. 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 bond 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 now turn to this objective.

2 Model and Estimation Strategy

Empirical studies of dynamic asset pricing models estimate the dynamics of a pricing kernel M_t that prices at time t an arbitrary payment Z_T at time T by $\mathbb{E}_t [(M_T / M_t) Z_T]$. Dynamic term structure models focus particular attention on pricing payoffs at different maturities T .

The dynamic term structure models we estimate fall within the broad class of models in which the pricing kernel is modelled as

$$dM_t = -M_t r(X_t) dt - M_t \Lambda(X_t)^\top dW_t$$

where X_t are latent factors with dynamics

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t.$$

The price P_t^T at time t of zero coupon bond³ that pays \$1 at time T is $\mathbb{E}_t[M_T / M_t]$ and depends critically on the dynamics of both the instantaneous short interest rate $r_t = r(X_t)$ and the market price of risk $\Lambda_t = \Lambda(X_t)$. A simple application of Itô's Lemma implies that zero coupon bond price dynamics follow

$$dP_t^T = \left[r_t P_t^T + \frac{\partial P_t^T}{\partial X_t} \cdot \sigma_t \Lambda_t \right] dt + \frac{\partial P_t^T}{\partial X_t} \cdot \sigma_t dW_t.$$

From (1) it is clear that expected excess returns of zero coupon bonds depend on both the market price of risk Λ_t as well as the volatility σ_t of the latent factors.

We estimate three 3-factor affine term structure models⁴ such that:

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

³In this paper we focus on modelling the swap rate and therefore the price of a zero coupon bond is the price of \$1 discounted at the relevant swap discount rate for that maturity.

⁴These models were introduced by ? and Dai and Singleton (2000). We use an extended affine market price of risk introduced by Cheridito et al. (2004) 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.

The drift is also affine under the risk neutral measure:

$$\mu_t^Q = K_0^Q + K_1^Q X_t$$

and the associated market price of risk is given by:

$$\Lambda(X_t) = (\sigma_t)^{-1} \left[K_0^P - K_0^Q + (K_1^P - K_1^Q) X_t \right]$$

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

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

In this affine setting, Duffie and Kan (1996) show that zero coupon bond prices are given by

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

where the functions A and B satisfy Riccati ODEs

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

We also include the prices of interest rate caps in our model estimation.

An interest rate cap is a financial derivative that caps the interest rate that is paid on the floating side of a swap. And so a cap is a portfolio of options on 3 month LIBOR. The price $C_t^N(\bar{C})$ of an N -period interest rate cap with strike rate \bar{C} and time Δt between floating interest payments is⁵

$$C_t^N(\bar{C}) = \sum_{n=2}^N \mathbb{E}_t \left[e^{-\int_t^{t+n\Delta t} r_r d\tau} \left(\frac{1}{P_{t+(n-1)\Delta t}^{t+n\Delta t}} - (1 + \bar{C} \Delta t) \right)^+ \right].$$

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, as we show in the appendix, when the solutions A and B to the Riccati ODEs are not known in closed form, numerical evaluation of the inverted Fourier transforms is computationally expensive for use in estimation. We use a more computationally efficient cumulant expansion technique to compute cap prices.⁶ The cumulant expansion method we develop is especially well-suited to option pricing in an affine framework and is described in more detail in Section C in the appendix.

Our data, obtained from Datastream, consists of Libor, swap rates, and at-the-money cap implied volatilities from January 1995 to March 2004. We use 3-month Libor and the entire term structure of swap rates to bootstrap swap zero rates at 1-, 2-, 3-, 5- and 10-years.⁷ Finally, we use at-the-money

⁵The market convention is that there is no cap payment for the first floating rate payment.

⁶Jarrow and Rudd (1982) were the first to use cumulant expansions in an asset pricing setting. Collin-Dufresne and Goldstein (2002) use cumulant expansions to compute swaption prices.

⁷Our bootstrap procedure assumes that forward swap zero rates are constant between

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.⁸ The full model specifications and estimation procedure 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 rate exactly and the remaining swap zero coupon rates are assumed to be priced with error.⁹ 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 used 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.¹⁰
3. Repeat these steps 1000 times to obtain a global maximum.

The parameter estimates are contained in Table 2.

observations.

⁸An $A_M(3)$ model has three latent factors with M factors having stochastic volatility.

⁹By 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.

¹⁰In an affine model, the conditional mean and variance are known in closed form as the solution to a linear constant coefficient ODE.

3 Cross Sectional Fit

Table 3 provides the root mean squared errors (in basis points) for the swap zero coupon rates. 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 $A_0(3)$ model has the lowest mean squared errors across term structure maturities. More importantly, the pricing errors are only slightly higher for the $A_1(3)^o$ and $A_2(3)^o$ models that are estimated with options than they are for the $A_1(3)$ and $A_2(3)$ models that are not estimated with options. Thus, including options in estimation does not appear to adversely affect the model's ability to successfully price the cross-section of interest rates.

Figure 1 plots at-the-money cap prices and Table 4 displays the root mean squared error in percentage terms for at-the-money caps with various maturities. While the $A_0(3)$ model had the lowest pricing errors for interest rates, it has the highest pricing errors for caps. The large cap pricing errors for the $A_0(3)$ model are due to its lack of factors with stochastic volatility. Since the $A_0(3)$ model does not contain stochastic volatility, we do not estimate it with options. The cap pricing errors for the $A_1(3)^o$ model are approximately half the size of the pricing errors for its $A_1(3)$ counterpart that is not estimated with options. More strikingly, the cap pricing errors for the $A_2(3)^o$ model are approximately one quarter the size of the pricing errors for the $A_2(3)$ model. Thus, while including options slightly increases the pricing errors for

the term structure of swap zero rates, it dramatically decreases the pricing errors for interest rate caps.¹¹

4 Matching Volatility

For the $A_1(3)^o$ and $A_2(3)^o$ models, cap prices are used in estimation and thus it is possible that these models are accurately capturing cap prices without accurately capturing interest rate volatility. As an additional measure of how well the models are capturing interest rate volatility, we also compute the prices of at-the-money swaptions. Swaptions differ from interest rate caps in that they are a single option on a long maturity swap rate rather than a portfolio of options on the 3-month Libor interest rate.

Figures 2 and 3 plot the times series of Black's swaption implied volatilities.¹² Tables 5 and 6 give the pricing errors for a cross section of swaption prices. The results for swaptions are similar to those for caps. The $A_0(3)$ model has the largest pricing errors. Again, the swaption pricing errors for the $A_1(3)^o$ and $A_2(3)^o$ models that are estimated with caps are significantly lower than their counterpart models $A_1(3)$ and $A_2(3)$ that are estimated without using options. Data from SwapPX indicates that typical bid-ask

¹¹It should be noted that none of the five models does a good job of pricing 1-year caps. 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 three-factor models because we are primarily interested in predicting changes in long term yields.

¹²We assume that the strike prices is the at-the-money forward swap rate implied by the model. This assumption is designed to minimize the effect of pricing errors in swap rates on the computation of swaption prices.

spreads are on the order of 2% implied volatility. Thus, the pricing errors for the $A_1(3)^o$ and $A_2(3)^o$ models are very close to the bid-ask spreads in these markets. Thus we find that these models are able to capture prices in both the bond and cap and swaptions markets (with the exception of the short end of the curve.)

In regards to pricing, these results differ somewhat from prior literature. Longstaff et al. (2001) and Han (2004) suggest that affine term structure models require a large number of parameters to simultaneously match both swaption and cap prices. Longstaff et al. (2001) price swaptions in estimation and find implied errors on cap prices of a similar magnitude to ours, but which under-price the caps on average whereas our model estimates have near zero average price errors for both cap and swaptions. Jagannathan et al. (2003) find that $A_N(N)$ models with independent factors do a very poor job of pricing caps whether or not they are included in the estimation. However, we use a more general price of risk and our computational procedure allows us to include affine models where cap prices are not known in closed form.

Our results are similar to Umantsev (2002) and Joslin (2005). Umantsev (2002) finds that low factor affine models can simultaneously price well both a cross section bonds and swaptions (though he does not consider caps.) Joslin (2005) finds the complementary result that including swaption prices in estimation gives models which price bonds, swaptions, and caps well.

Implied volatilities from caps and swaptions are forward looking and, in the case of stochastic volatility models, also contain risk premia. The realized

volatility however is not observed. For estimates of conditional volatility based on historical data we use a 26 week rolling window, an exponential weighted moving average (EWMA) with a 26-week half-life, and estimate an EGARCH(1,1) for each maturity.

Figure 4 plots the model's conditional volatility of zero coupon rates against these estimates of conditional volatility using historical data. None of the models do a good job of tracking the various estimates of the volatility of the 6-month zero coupon rate, though the $A_1(3)^o$ and $A_2(3)^o$ at least appear to get the level right.¹³ However, for the 2- and 5- year maturities, the conditional volatility of the $A_1(3)^o$ and $A_2(3)^o$ models more closely tracks the various estimates of conditional volatility. The $A_2(3)$ model completely misses the level of volatility for the 6-month and 2-year zero coupon rates. Though, on average, the $A_2(3)$ model matches the level of the volatility of the 5-year zero coupon rate, it appears to miss the dynamics. The $A_1(3)$ model does a better job than the $A_2(3)$ model at matching the various historical estimates of conditional volatility. However, in each case, the $A_1(3)$ and $A_2(3)$ models are worse than their $A_1(3)^o$ and $A_2(3)^o$ counterparts.

¹³As noted earlier, Dai and Singleton (2002) suggest that a fourth factor is required to capture the dynamics of the short end of the yield curve. Collin-Dufresne et al. (2004) are able to match the volatility of the short end with an unspanned stochastic volatility model.

5 Predictability of Excess Returns

Table 7 presents evidence on the predictability of excess returns for long term interest rates for the in-sample period from January 1995 to March 2004. R^2 's are calculated as

$$R^2 = 1 - \text{var}(R_{t,n}^{\text{expected}} - R_{t,t+1}^n) / \text{var}(R_{t,t+1}^n),$$

where $\text{var}(\cdot)$ denotes variance, $R_{t,n}^{\text{expected}}$ are weekly model implied expected returns for discount bonds with n years to maturity, and $R_{t,t+1}^n$ are weekly realized returns for the corresponding bond. We include R^2 's for each model we estimated, as well as R^2 's from three versions of the regressions of excess returns on forward rates as performed in Cochrane and Piazzesi (2005).¹⁴

On average, amongst models that were estimated without options, the $A_0(3)$ model has higher excess return predictability than the $A_1(3)$ model, which in turn has higher predictability than the $A_2(3)$ model. Both Duffee (2002) and Dai and Singleton (2002) also estimate three-factor term structure

¹⁴For different maturities, Cochrane and Piazzesi (2005) run regressions of yields variations on a linear combination of forward rates. Letting p_t^n and y_t^n denote respectively the price and yield to maturity of a n -year discount bond at time t , for each fixed n they regress:

$$r_{t+1}^n - y_t^1 = \beta_0^n + \beta_1^n y_t^1 + \beta_2^n f_t^2 + \beta_3^n f_t^3 + \beta_4^n f_t^4 + \beta_5^n f_t^5 + \epsilon_{t+1}^n,$$

where r_{t+1}^n is the holding period return from buying an n period discount bond at time t and selling it at time $t + 1$, and $f_t^i = p_t^{i-1} - p_t^i, i = 2, \dots, 5$ is the time t one period forward rate for loans between the maturities $i - 1$ and i . 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, $CP_{5,10}$ use only 5 one year forward rates (which begin in 0,2,4,6, and 8 years) as regressors.

models without options and find that the $A_0(3)$ model has the best performance in terms of predictability. When options are included in estimation, the predictability of both the $A_1(3)^o$ and $A_2(3)^o$ models improve dramatically over their $A_1(3)$ and $A_2(3)$ counterparts. On average, the R^2 's for the $A_1(3)^o$ model are two to three times as large as those for the $A_0(3)$. The difference is dramatic for the 10-year maturity where the R^2 for the $A_0(3)$ model is only 2.5% but the R^2 for the $A_1(3)^o$ is 33.1%.

Moreover, the R^2 's are much closer in magnitude to those obtained from the regressions in Cochrane and Piazzesi (2005). The regressions in Cochrane and Piazzesi (2005) are designed to only match excess returns and so they serve as somewhat of an upper bound for the the level of predictability of excess returns.

Table 8 provides R^2 's for the out-of-sample period from April 1988 to December 1994. (Recall that the models were estimated with historical data from January 1995 to March 2004, which corresponded to the availability of cap data in Datastream.) The $A_0(3)$ and $A_2(3)$ models do extremely poorly out-of-sample, while CP_{10} seems to be overfitting in-sample data (which motivating including the $CP_{5,10}$). As was the case with in-sample predictability, the inclusion of options in the $A_1(3)^o$ and $A_2(3)^o$ models dramatically improves their out-of-sample predictability. Equally as striking, the out-of-sample predictability of the $A_1(3)^o$ model estimated with options is on par with that of the CP_5 and CP_{10} results from the regressions in Cochrane and Piazzesi (2005).

Figure 5 plots the realized excess returns as well as the expected excess returns for the $A_0(3)$, $A_1(3)$, and $A_1(3)^o$ models and the CP_5 regressions. The variation in expected excess returns is higher for the $A_1(3)^o$ and $A_2(3)^o$ models than for their $A_1(3)$ and $A_2(3)$ counterparts, presumably because these models capture more time variation in volatility when they are estimated with options. The results in Table 9 confirm this observation. In addition, not only is the level of predictability of excess returns higher for the $A_1(3)^o$ and $A_2(3)^o$ models than for the $A_0(3)$ model, the variation in the predict excess returns is actually lower. Since there is no time variation in volatility for the $A_0(3)$ model, all of the variation in expected excess returns is due to variation in the market price of risk. Thus, the $A_0(3)$ model appears to overstate the true amount of variation in the market prices of risk. The variation in expected excess returns for the $A_1(3)^o$ and $A_2(3)^o$ models is also lower than that for the CP_5 regressions. However, the CP_5 regression is not an economic model and therefore the expected excess returns cannot be decomposed into volatility and the market prices of risk.

6 Conclusion

We estimate three-factor affine term structure models jointly on both swap rates and interest rate cap prices. When we incorporate information in interest rate caps, we significantly improve the model's ability to price swaptions and match realized volatility without impairing its ability to capture the term

structure of interest rates. Furthermore, the model's that are estimated with options are dramatically better at predicting excess returns for long-term swaps over short-term swaps, both in- and out-of-sample. In contrast to previous literature, the arbitrage-free models with the most predictive power contain a stochastic volatility component. Our results indicate that interest rate options contain valuable information about term structure dynamics that cannot be extracted from interest rates alone.

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.

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 a ne models: Theory and evidence. Working Paper, Princeton University, 15 October 2004.
- John H. Cochrane and Monika Piazzesi. Bond risk premia. *American Economic Review forthcoming*, 2005.
- Pierre Collin-Dufresne and Robert S. Goldstein. Pricing swaptions within an affine framework. *Journal of Derivatives*, 10(1):1–18, Fall 2002.
- Pierre Collin-Dufresne, Robert S. Goldstein, and Christopher S. Jones. Can the volatility of interest rates be extracted from the cross section of bond yields? an investigation of unspanned stochastic volatility. Working Paper, 5 August 2004.
- 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.
- Ravi Jagannathan, Andrew Kaplin, and Steve Sun. An evaluation of multi-factor cir models using libor, swap rates, and cap and swaption prices. *Journal of Econometrics*, 116(1-2):113–146, September-October 2003.
- Robert Jarrow and Andrew Rudd. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369, November 1982.
- Scott Joslin. Completeness in fixed income markets. 2005. Working Paper, Stanford University.
- 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

The short rate is given by $r_t = \rho_0 + \rho_1 \cdot X_t$, where X_t is a Markov state variable with dynamics and the physical and risk neutral measures given by:

$$\begin{aligned}dX_t &= (K_0^P + K_1^P X_t) + \sigma_t dB_t^P \\dX_t &= (K_0^Q + K_1^Q X_t) + \sigma_t dB_t^Q\end{aligned}$$

and where the conditional variance is affine in the state: $\sigma_t \sigma_t' = H_0 + H_1 \cdot X_t$.

In the $A_0(3)$ model, $H_1 \equiv 0$, so none of the three factors in X_t have stochastic volatility. In the $A_1(3)$ model, one of the factors in X_t drives stochastic volatility, and in the $A_2(3)$ model, two of the factors in X_t drive stochastic volatility. For each model, Dai and Singleton (2000) and Cheridito et al. (2004) identify the necessary restrictions required to ensure that the stochastic processes are admissible, 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.

A_0 (3) Model Specification

$$\begin{aligned}
 K_1^P &= \begin{bmatrix} K_{1,11}^P & 0 & 0 \\ K_{1,21}^P & K_{1,22}^P & 0 \\ K_{1,31}^P & K_{1,32}^P & K_{1,33}^P \end{bmatrix} & K_0^P &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
 \sigma_t &= I_3 \\
 \rho_{1,1} &\geq 0, \quad \rho_{1,2} \geq 0, \quad \rho_{1,3} \geq 0
 \end{aligned}$$

A_1 (3) Model Specification

$$\begin{aligned}
 K_1^P &= \begin{bmatrix} K_{1,11}^P & 0 & 0 \\ K_{1,21}^P & K_{1,22}^P & K_{1,23}^P \\ K_{1,31}^P & K_{1,32}^P & K_{1,33}^P \end{bmatrix} & K_0^P &= \begin{bmatrix} K_{0,1}^P \\ 0 \\ 0 \end{bmatrix} \\
 K_1^Q &= \begin{bmatrix} K_{1,11}^Q & 0 & 0 \\ K_{1,21}^Q & K_{1,22}^Q & K_{1,23}^Q \\ K_{1,31}^Q & K_{1,32}^Q & K_{1,33}^Q \end{bmatrix} & K_0^Q &= \begin{bmatrix} K_{0,1}^Q \\ 0 \\ 0 \end{bmatrix} \\
 \sigma_t &= \begin{bmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{12} X_t^1} & 0 \\ 0 & 0 & \sqrt{1 + \beta_{13} X_t^1} \end{bmatrix} \\
 K_{0,1}^P &\geq \frac{1}{2}, \quad K_{0,1}^Q \geq \frac{1}{2} \\
 \beta_{12} &\geq 0, \quad \beta_{13} \geq 0 \\
 \rho_{1,2} &\geq 0, \quad \rho_{1,3} \geq 0
 \end{aligned}$$

A₂ (3) Model Specification

$$\begin{aligned}
 K_1^P &= \begin{bmatrix} K_{1,11}^P & K_{1,12}^P & 0 \\ K_{1,21}^P & K_{1,22}^P & 0 \\ K_{1,31}^P & K_{1,32}^P & K_{1,33}^P \end{bmatrix} & K_0^P &= \begin{bmatrix} K_{0,1}^P \\ K_{0,2}^P \\ 0 \end{bmatrix} \\
 K_1^Q &= \begin{bmatrix} K_{1,11}^Q & K_{1,12}^Q & 0 \\ K_{1,21}^Q & K_{1,22}^Q & 0 \\ K_{1,31}^Q & K_{1,32}^Q & K_{1,33}^Q \end{bmatrix} & K_0^Q &= \begin{bmatrix} K_{0,1}^Q \\ K_{0,2}^Q \\ 0 \end{bmatrix} \\
 \sigma_t &= \begin{bmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{1 + \beta_{13}X_t^1 + \beta_{23}X_t^2} \end{bmatrix} \\
 K_{0,1}^P &\geq \frac{1}{2}, & K_{0,2}^P &\geq \frac{1}{2}, & K_{0,1}^Q &\geq \frac{1}{2}, & K_{0,2}^Q &\geq \frac{1}{2} \\
 K_{1,12}^P &\geq 0, & K_{1,21}^P &\geq 0, & K_{1,12}^Q &\geq 0, & K_{1,21}^Q &\geq 0 \\
 \beta_{13} &\geq 0, & \beta_{23} &\geq 0 \\
 \rho_{1,3} &\geq 0
 \end{aligned}$$

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 i.i.d. 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 likelihood 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 slightly 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 $(M, \text{vec}(V))$, 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 $\Delta_{i,j} = \text{vec}(H_{1,\cdot,i})_j$. Rather than con-

sidering 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 Cap Valuation via a Cumulant Expansion

Recall that P_t^T is the price at time t of \$1 paid at time T . The price $C_t^N(\bar{C})$ of an N -period interest rate cap with strike rate \bar{C} and time Δt between floating interest payments is

$$\begin{aligned} C_t^N(\bar{C}) &= \sum_{n=2}^N \mathbb{E}_t \left[\frac{M_{t+n\Delta t}}{M_t} \left(\frac{1}{P_{t+(n-1)\Delta t}^{t+n\Delta t}} - (1 + \bar{C}\Delta t) \right)^+ \right] \\ &= e^{-A(\Delta t)} \sum_{n=2}^N G_t(-A(\Delta t) - \ln(1 + \bar{C}\Delta t); -B(\Delta t), B(\Delta t), (n-1)\Delta t) \\ &\quad - e^{-A(\Delta t)} (1 + \bar{C}\Delta t) \sum_{n=2}^N G_t(-A(\Delta t) - \ln(1 + \bar{C}\Delta t); 0, B(\Delta t), (n-1)\Delta t), \end{aligned}$$

where

$$G_t(y; b, \gamma, \tau) := \mathbb{E}_t \left[\frac{M_{t+\tau}}{M_t} e^{b^\top X_{t+\tau}} \{ \gamma^\top X_{t+\tau} \leq y \} \right].$$

Thus, cap valuation requires that we be able to efficiently compute G_t .

By the Lévy inversion formula,

$$G_t(y; b, \gamma, \tau) = \frac{1}{2} \hat{G}_t(0; b, \gamma, \tau) - \frac{1}{\pi} \int_0^\infty \frac{1}{v} \operatorname{Im} \left[e^{-ivy} \hat{G}_t(v; b, \gamma, \tau) \right] dv,$$

where \hat{G}_t is the Fourier transform of G_t . In an affine framework \hat{G}_t is given by

$$\begin{aligned} \hat{G}_t(v; b, \gamma, \tau) &= \mathbb{E}_t \left[\frac{M_{t+\tau}}{M_t} e^{(b+iv\gamma)^\top X_{t+\tau}} \right] \\ &= e^{A(b+iv\gamma, \tau) + B(b+iv\gamma, \tau)^\top X_t}, \end{aligned}$$

where, A and B satisfy the Riccati ODEs

$$\begin{aligned} \frac{\partial B(b+iv\gamma, u)}{\partial u} &= -\rho_1 + \mathcal{K}_1^{\mathbb{Q}\top} B(b+iv\gamma, u) + \frac{1}{2} \beta \Delta [B(b+iv\gamma, u)] B(b+iv\gamma, u), \\ \frac{\partial A(b+iv\gamma, u)}{\partial u} &= -\rho_0 + \mathcal{K}_0^{\mathbb{Q}\top} B(b+iv\gamma, u) + \frac{1}{2} \alpha^\top \Delta [B(b+iv\gamma, u)] B(b+iv\gamma, u), \end{aligned}$$

with boundary conditions

$$\begin{aligned} B(b+iv\gamma, 0) &= b+iv\gamma, \\ A(b+iv\gamma, 0) &= 0. \end{aligned}$$

If the affine model is such that the solutions A and B to the Riccati ODEs are known in closed form, then cap valuation only requires numerical evaluation of a 1-dimensional integral. However, in the general case, the Riccati ODEs must be solved numerically and thus valuing a cap using the

Lévy inversion formula is not computationally feasible for model estimation.

Instead, we use a more computationally efficient cumulant expansion technique to compute cap prices. The cumulant expansion requires that we compute the Taylor series expansion of the log of the Fourier transform of G_t . Define the cumulants c_m by

$$\begin{aligned} c_m &:= \frac{\partial^m \ln \hat{G}_t(0; b, \gamma, \tau)}{\partial (iv)^m} \\ &= i^{3m} \left\{ \underbrace{\frac{\partial^m A(b + iv\gamma, \tau)}{\partial v^m}}_{\partial_v^m A(\tau)} + \underbrace{\frac{\partial^m B(b + iv\gamma, \tau)^\top}{\partial v^m}}_{\partial_v^m B(\tau)} X_t \right\} \Big|_{v=0}, \end{aligned}$$

so that

$$\ln \hat{G}_t(v; b, \gamma, \tau) = \ln \hat{G}_t(0; b, \gamma, \tau) + \sum_{m=1}^{\infty} \frac{1}{m!} c_m (iv)^m.$$

In an affine framework, the cumulants are affine in the state vector X_t with coefficients that again satisfy Riccati ODEs,

$$\begin{aligned} \partial_v^1 B(u) &= \mathcal{K}_1^{\mathbb{Q}\top} \partial_v^1 B(u) + \beta \Delta \left[B(u)^\top \Sigma \right] \Sigma^\top \partial_v^1 B(u), \quad \partial_v^1 B(0) = i\gamma, \\ \partial_v^1 A(u) &= \mathcal{K}_0^{\mathbb{Q}\top} \partial_v^1 B(u) + \alpha^\top \Delta \left[B(u)^\top \Sigma \right] \Sigma^\top \partial_v^1 B(u), \quad \partial_v^1 A(0) = 0, \end{aligned}$$

and for $m > 1$,

$$\begin{aligned}\partial_v^m B(u) &= \mathcal{K}_1^{\mathbb{Q}\top} \partial_v^m B(u) + \sum_{k=0}^m \varphi_k^m \beta \Delta \left[\partial_v^{m-k} B(u)^\top \Sigma \right] \Sigma^\top \partial_v^k B(u), \quad \partial_v^m B(0) = 0, \\ \partial_v^m A(u) &= \mathcal{K}_0^{\mathbb{Q}\top} \partial_v^m B(u) + \sum_{k=0}^m \varphi_k^m \alpha^\top \Delta \left[\partial_v^{m-k} B(u)^\top \Sigma \right] \Sigma^\top \partial_v^k B(u), \quad \partial_v^m A(0) = 0,\end{aligned}$$

where $\varphi_k^m = \binom{m}{k}$ if $m \neq 2k$ and $\varphi_k^m = \frac{1}{2} \binom{m}{k}$ if $m = 2k$.

Once we have computed the cumulants, we can accurately approximate G_t by

$$G_t(y; b, \gamma, \tau) \approx \sum_{m=0}^M [\chi_{-1}^m \Phi_{-1}(y - c_1) + \chi_0^m \Phi_0(y - c_1)]$$

where

$$\begin{aligned}\Phi_{-1}(y) &= \frac{1}{\sqrt{2\pi c_2}} e^{-\frac{y^2}{2c_2}} \\ \Phi_0(y) &= \int_{-\infty}^y \Phi_{-1}(z) dz,\end{aligned}$$

and the coefficients χ_{-1}^m and χ_0^m are related to the cumulants as described below. Φ_{-1} and Φ_0 are just the density and cumulative distribution of the Normal distribution. There exist accurate approximations to the cumulative Normal density, therefore computation of cap prices using a cumulant expansion does not require any numerical integration (aside from solving Riccati ODEs). We now turn to determining the coefficients χ_{-1}^m and χ_0^m .

Define a_m to be the coefficients in a Taylor series expansion of

$$\hat{G}_t(v; b, \gamma, \tau) e^{-[c_1(iv) + \frac{1}{2}c_2(iv)^2]},$$

about $v = 0$, so that

$$\hat{G}_t(v; b, \gamma, \tau) = e^{c_1(iv) - \frac{1}{2}c_2 v^2} \sum_{m=0}^{\infty} a_m v^m.$$

Then

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izv} \hat{G}_t(v; b, \gamma, \tau) dv &= \sum_{m=0}^{\infty} a_m \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(z-c_1)v - \frac{1}{2}c_2 v^2} v^m dv \\ &= \sum_{m=0}^{\infty} a_m \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial^m e^{uv - \frac{1}{2}c_2 v^2}}{\partial u^m} \Big|_{u=-i(z-c_1)} dv \\ &= \sum_{m=0}^{\infty} \frac{\partial^m}{\partial u^m} \left\{ a_m \frac{1}{\sqrt{2\pi c_2}} e^{\frac{u^2}{2c_2}} \right\} \Big|_{u=-i(z-c_1)} \\ &\approx \sum_{m=0}^M \frac{\partial^m}{\partial u^m} \left\{ a_m \frac{1}{\sqrt{2\pi c_2}} e^{\frac{u^2}{2c_2}} \right\} \Big|_{u=-i(z-c_1)} \\ &=: \frac{1}{\sqrt{2\pi c_2}} e^{\frac{-(z-c_1)^2}{2c_2}} \sum_{m=0}^M \lambda_m (z - c_1)^m, \end{aligned}$$

where the last line defines the coefficients λ_m .

Then by the inverse Fourier transform,

$$\begin{aligned}
G_t^d(y; b, \gamma, \tau) &= \int_{-\infty}^y \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izv} \hat{G}_t^d(v; b, \gamma, \tau) dv dz \\
&\approx \sum_{m=0}^M \lambda_m \underbrace{\int_{-\infty}^y \frac{1}{\sqrt{2\pi c_2}} e^{-\frac{(z-c_1)^2}{2c_2}} (z-c_1)^m dz}_{\Phi_m(y-c_1)},
\end{aligned}$$

$\Phi_m(y)$ can be expressed in terms of $\Phi_{-1}(y)$ and $\Phi_0(y)$ via the recursive relationship

$$\begin{aligned}
\Phi_{-1}(y) &= \frac{1}{\sqrt{2\pi c_2}} e^{-\frac{y^2}{2c_2}} \\
\Phi_0(y) &= \int_{-\infty}^y \Phi_{-1}(z) dz, \\
\Phi_m(y) &= -c_2 \int_{-\infty}^y z^{m-1} d\Phi_{-1}(z) \\
&= -c_2 [y^{m-1}\Phi_{-1}(y) - (m-1)\Phi_{m-2}(y)].
\end{aligned}$$

Therefore, $G_t(y; b, \gamma, \tau)$ is of the form

$$G_t(y; b, \gamma, \tau) \approx \sum_{m=0}^M [\chi_{-1}^m \Phi_{-1}(y-c_1) + \chi_0^m \Phi_0(y-c_1)],$$

as desired.

Finally, M must be chosen to balance accuracy and computational speed. We follow Collin-Dufresne and Goldstein (2002) and choose $M = 7$ in our estimations.

D Tables and Figures

	2 Year	3 Year	4 Year	5 Year
Slope only	3.4%	6.3%	7.9%	8.7%
Slope and cap implied volatility	15.1%	25.2%	31.7%	35.2%
All yields	38.1%	45.3%	50.8%	54.7%

Table 1: Regression of Excess Returns.

This table shows the R^2 from regressions of (overlapping) one year excess returns of 2-year to 5-year bonds for various regressors. The sample period is June 1995 to March 2004.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
$K_{1,1}^{0P}$	0	3.607 (1.55)	2.264 (1.776)	0.5711 (2.298)	0.7991 (3.648)
$K_{2,1}^{0P}$	0	0	0	1.319 (3.842)	0.9439 (2.285)
$K_{3,1}^{0P}$	0	0	0	0	0
$K_{1,1}^{0Q}$	1.386 (4.33)	1.097 (0.1348)	1.301 (0.3113)	1.919 (0.437)	1.376 (0.1462)
$K_{2,1}^{0Q}$	0.4015 (2.877)	0.8248 (0.3585)	4.263 (1.174)	0.8986 (0.255)	0.7215 (0.5222)
$K_{3,1}^{0Q}$	-0.2268 (0.2842)	-0.2536 (4.76)	2.54 (1.011)	-1.569 (0.8322)	0.3154 (1.16)
$K_{1,1}^{1P}$	-0.02769 (0.1655)	-1.534 (0.6193)	-5.094e-005 (0.4953)	-0.7572 (0.4223)	-0.9914 (1.207)
$K_{1,2}^{1P}$	0	0	0	1.025 (0.5063)	0.4955 (2.057)
$K_{1,3}^{1P}$	0	0	0	0	0
$K_{2,1}^{1P}$	0.664 (0.3809)	0.4467 (0.5795)	-0.6818 (0.7408)	0.4932 (0.819)	0.5848 (0.8917)
$K_{2,2}^{1P}$	-0.3399 (0.1812)	-0.592 (0.6694)	-1.046 (0.3693)	-1.031 (1.12)	-1.45 (1.005)
$K_{2,3}^{1P}$	0	0.006785 (0.5246)	-1.758 (1.013)	0	0
$K_{3,1}^{1P}$	-0.9474 (0.3542)	-0.8855 (1.006)	-0.4861 (0.4114)	-1.284 (0.2839)	-0.6992 (0.8984)
$K_{3,2}^{1P}$	-0.4975 (0.3785)	-0.5759 (0.7353)	-0.6248 (0.2242)	1.67 (0.3843)	-1.168 (1.485)
$K_{3,3}^{1P}$	-1.181 (0.5553)	-0.5555 (0.6346)	-1.414 (0.5087)	-0.1309 (0.07153)	-0.03759 (0.06582)
$K_{1,1}^{1Q}$	-1.153 (0.09701)	-0.5376 (0.009432)	-0.5307 (0.01271)	-1.613 (0.03691)	-0.5734 (0.05517)
$K_{1,2}^{1Q}$	1.783 (0.299)	0	0	1.18 (0.04536)	0 (0.0674)
$K_{1,3}^{1Q}$	1.597 (0.03679)	0	0	0	0
$K_{2,1}^{1Q}$	0.1279 (0.01355)	-0.5693 (0.03321)	-2.213 (0.4629)	1.023 (0.04626)	1.433 (0.1664)
$K_{2,2}^{1Q}$	-0.4049 (0.05457)	-0.3371 (0.03703)	-1.003 (0.07605)	-1.425 (0.04184)	-2.376 (0.09309)
$K_{2,3}^{1Q}$	-0.4135 (0.009265)	-0.1229 (0.01658)	-1.916 (0.2753)	0	0
$K_{3,1}^{1Q}$	-0.1348 (0.007076)	-0.7597 (0.1177)	-0.926 (0.1237)	-1.618 (0.03916)	0.579 (0.1147)
$K_{3,2}^{1Q}$	-0.1364 (0.009808)	-2.388 (0.1048)	-0.6465 (0.09368)	2.992 (0.04763)	-2.122 (0.2817)
$K_{3,3}^{1Q}$	-0.2887 (0.02507)	-1.494 (0.09641)	-1.34 (0.1018)	-0.133 (0.00154)	-0.04871 (0.001681)
$\alpha_{1,1}$	1	0	0	0	0
$\alpha_{2,1}$	1	1	1	0	0
$\alpha_{3,1}$	1	1	1	1	1
$\beta_{1,1}$	0	1	1	1	1
$\beta_{1,2}$	0	0.3849 (0.03349)	3.363 (1.715)	0	0
$\beta_{1,3}$	0	2.581e-008 (0.03023)	0 (0.04901)	0.03079 (0.02461)	0.5111 (0.4016)
$\beta_{2,1}$	0	0	0	0	0
$\beta_{2,2}$	0	0	0	1	1
$\beta_{2,3}$	0	0	0	0.0002741 (0.02945)	0.438 (0.3572)
$\beta_{3,1}$	0	0	0	0	0
$\beta_{3,2}$	0	0	0	0	0
$\beta_{3,3}$	0	0	0	0	0
ρ^0	-0.1934 (0.9732)	0.07265 (0.0314)	0.001041 (0.1396)	0.01166 (0.01518)	0.2885 (0.1359)
$\rho_{1,1}^1$	0.01278 (0.0005886)	0.0002245 (0.0003265)	0.0001308 (0.0003784)	0.001844 (7.342e-005)	0.01011 (0.002518)
$\rho_{2,1}^1$	0.008455 (0.001243)	0.0014 (0.0003625)	0.0007724 (0.0002196)	-0.003853 (0.0001577)	0.01435 (0.002262)
$\rho_{3,1}^1$	0 (0.001307)	0.008652 (0.0003321)	0.01049 (0.001036)	0.004448 (8.368e-005)	0.005957 (0.00119)
M	0	1	1	2	2
N	3	3	3	3	3
$timestep$	0.01923	0.01923	0.01923	0.01923	0.01923
$LogLikelihood$	40.17	71.44	40.6	72.17	39.55

Table 2: Parameter Estimates.

This table presents all parameter values for the different affine term structure models estimated. 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 caps 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. (2004).

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
3 Month	0.0	0.0	0.0	0.0	0.0
1 Year	13.4	13.3	13.9	14.0	14.2
2 Year	0.0	0.0	0.0	0.0	0.0
3 Year	4.3	5.5	4.3	5.7	4.4
5 Year	5.3	8.0	5.5	8.2	5.5
7 Year	3.8	6.4	4.2	6.6	4.2
10 Year	0.0	0.0	0.0	0.0	0.0

Table 3: Relative Pricing Errors in % 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 caps were measured with error.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
1 Year	202.9	67.8	80.1	44.1	272.5
2 Year	73.8	17.7	24.4	17.1	90.1
3 Year	54.3	11.7	21.1	11.0	57.1
4 Year	45.9	9.5	21.4	8.7	43.1
5 Year	40.4	8.8	22.0	8.0	36.1
7 Year	34.4	8.3	22.5	7.5	30.4
10 Year	29.2	9.3	23.9	8.4	26.8

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 caps were measured with error.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
In 1 yr-For 1 yr	66.1	14.2	24.9	21.3	57.1
In 1 yr-For 2 yr	51.7	10.9	28.3	14.3	33.7
In 1 yr-For 3 yr	40.9	10.2	29.6	10.3	25.3
In 1 yr-For 4 yr	32.4	9.9	29.7	8.8	24.5
In 1 yr-For 5 yr	26.3	9.5	29.5	8.6	25.5
In 3 months-For 1 yr	95.1	24.0	29.7	35.4	108.5
In 3 months-For 2 yr	62.1	16.6	29.9	25.3	53.9
In 3 months-For 3 yr	48.6	13.3	32.1	18.1	32.6
In 3 months-For 4 yr	37.8	11.9	33.0	14.5	25.2
In 3 months-For 5 yr	30.1	11.5	33.0	13.0	24.7

Table 5: Relative Pricing Errors in % for At-the-Money Swaption
This table shows the root mean square relative pricing errors in % for at-the-money swaptions. 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 swaptions were not including in the estimation in any of the models.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
In 1 yr-For 1 yr	16.0	4.2	5.8	6.5	9.9
In 1 yr-For 2 yr	10.2	2.8	6.4	4.1	7.5
In 1 yr-For 3 yr	7.3	2.4	6.5	2.9	7.3
In 1 yr-For 4 yr	5.5	2.3	6.3	2.5	7.6
In 1 yr-For 5 yr	4.6	2.1	6.1	2.4	7.8
In 3 months-For 1 yr	33.1	10.0	13.9	9.3	24.7
In 3 months-For 2 yr	16.2	5.1	6.5	7.1	10.2
In 3 months-For 3 yr	9.8	3.6	6.6	5.0	8.4
In 3 months-For 4 yr	6.7	3.1	7.0	4.3	8.7
In 3 months-For 5 yr	5.3	3.0	7.2	4.3	9.3

Table 6: At-the-Money Swaption Implied Volatility Errors

This table shows the root mean square implied volatility errors for at-the-money swaptions. 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 swaptions were not including in the estimation in any of the models.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$	CP_5	CP_{10}	$CP_{5,10}$
2 Yr	-36.4	10.4	1.4	-4.0	6.3	27.9	41.5	34.7
3 Yr	4.3	24.7	4.8	6.3	6.7	37.3	49.2	43.2
4 Yr	14.1	27.7	7.9	11.2	2.6	43.5	54.1	48.6
5 Yr	15.8	29.4	9.2	13.1	0.1	47.4	57.0	51.9
6 Yr	14.8	30.6	9.6	13.8	-1.6	0.0	58.6	53.6
7 Yr	12.5	31.7	9.7	14.0	-2.9	0.0	59.8	54.9
8 Yr	10.0	31.8	9.5	13.7	-3.8	0.0	60.2	55.6
9 Yr	6.5	32.6	9.5	13.6	-4.6	0.0	60.5	55.8
10 Yr	2.5	33.1	9.4	13.3	-5.2	0.0	60.8	56.1

Table 7: In-Sample Predictability of Excess Returns (R^2 's in %)

This Table presents R^2 s obtained from projections of weekly realized zero coupon returns, for different maturities, on model in-sample 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. $CP_{5,10}$ use 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 caps were measured with error.

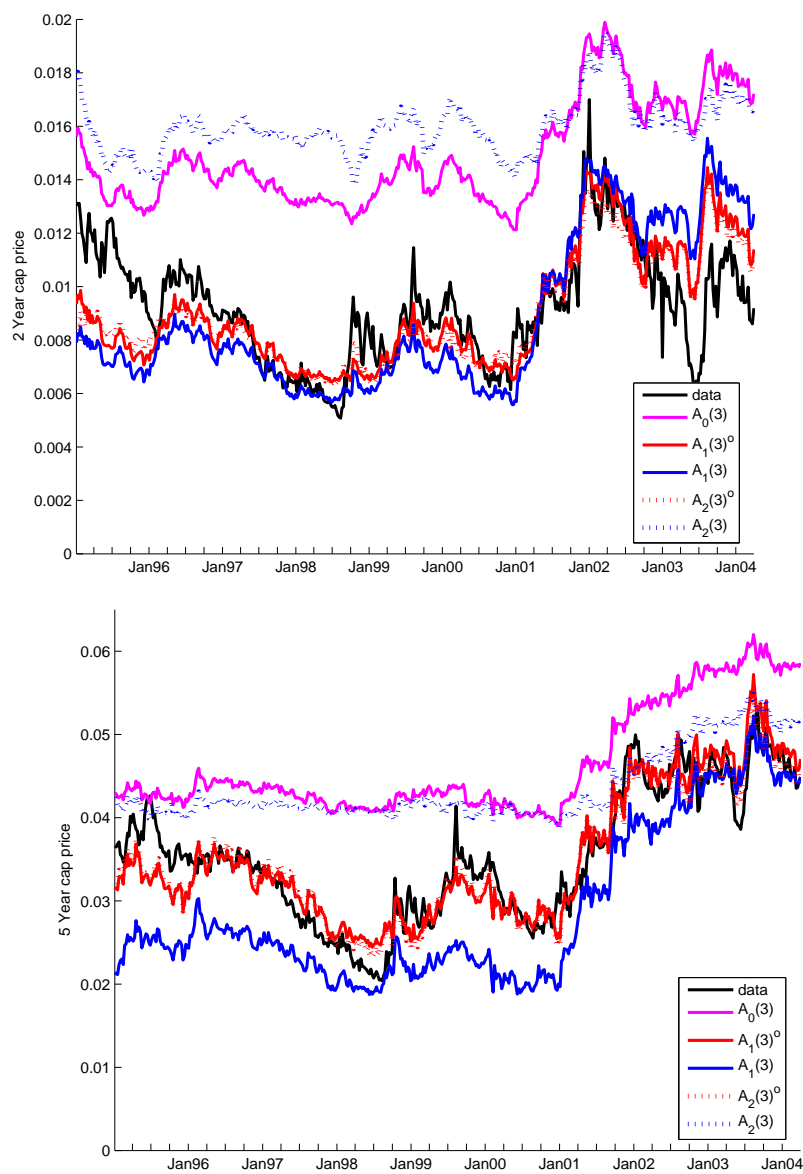


Figure 1: Cap Prices

The top figure plots 2-year at-the-money cap prices. The bottom figure plots 5-year at-the-money cap prices. The actual prices are plotted with a solid black line. The prices from the $A_0(3)$ model plotted with a solid pink line. The prices from the $A_1(3)$ model are plotted with a solid blue line and the prices from the $A_1(3)^\circ$ model are plotted with a solid red line. The prices from the $A_2(3)$ model are plotted with a dashed blue line and the prices from the $A_2(3)^\circ$ model are plotted with a dashed red line. The $A_0(3)$, $A_1(3)$, and $A_2(3)$ models were estimated by inverting 3-, 2-year, and 10-year swap zeros and measuring 1-, 3-, 5-, and 7-year zeros with error. The $A_1(3)^\circ$ and $A_2(3)^\circ$ models were estimated with the additional assumption that 1-, 2-, 3-, 4-, 5-, 7-, and 10-year at-the-money caps were measured with error.

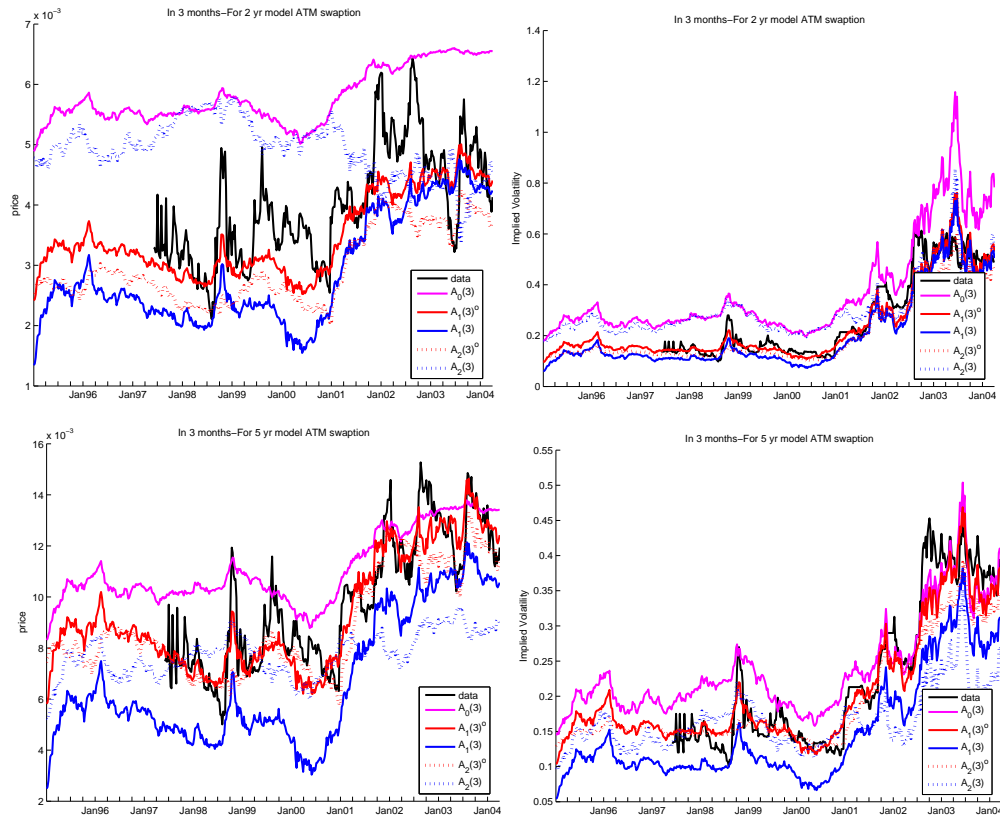


Figure 2: In 3 Months At-the-Money Swaption Implied Volatilities
 These figures plot prices and Black's implied volatilities for at-the-money in-3-months-for-2-year and in-3-months-for-5-year swaptions. The at-the-money strike rates are the forward swap rates which are taken from the model. The data are plotted with a solid black line. The values from the $A_0(3)$ model plotted with a solid pink line. The values from the $A_1(3)$ model are plotted with a solid blue line and the values from the $A_1(3)^o$ model are plotted with a solid red line. The values from the $A_2(3)$ model are plotted with a dashed blue line and the values from the $A_2(3)^o$ model are plotted with a dashed red line. 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.

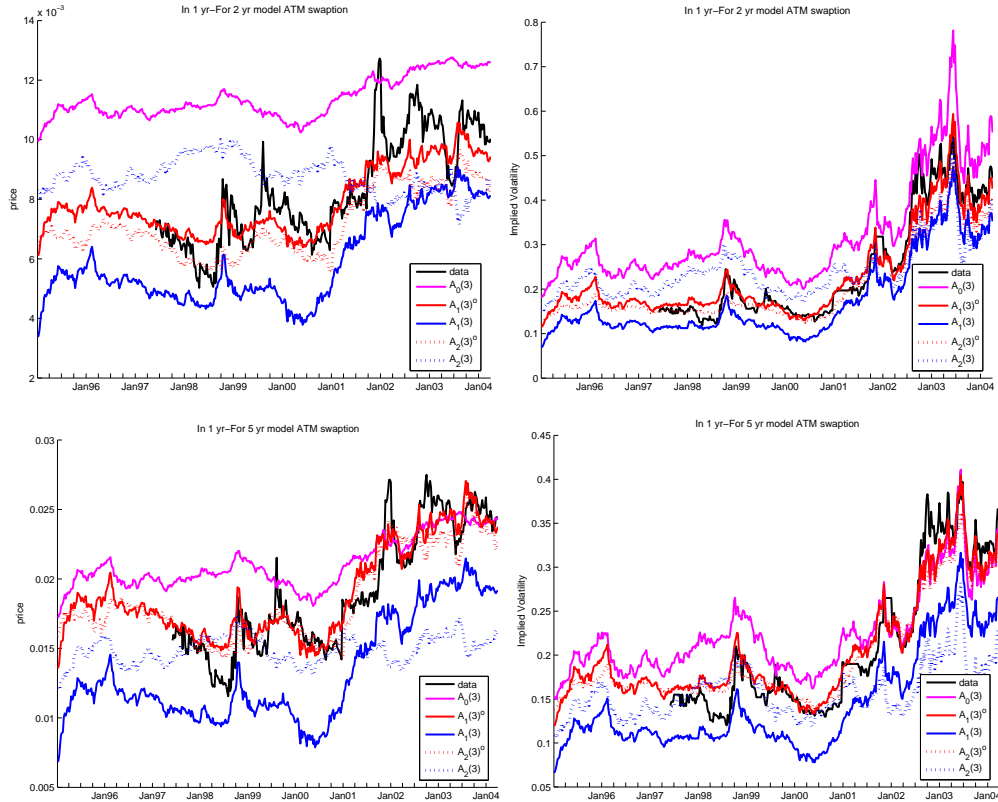


Figure 3: In 1 Year At-the-Money Swaption Implied Volatilities

These figures plot prices and Black's implied volatilities for at-the-money in-1-year-for-2-year and in-1-year-for-5-year swaptions. The at-the-money strike rates are the forward swap rates which are taken from the model. The data are plotted with a solid black line. The values from the $A_0(3)$ model plotted with a solid pink line and the values from the $A_1(3)$ model are plotted with a solid blue line and the values from the $A_1(3)^o$ model are plotted with a solid red line. The values from the $A_2(3)$ model are plotted with a dashed blue line and the values from the $A_2(3)^o$ model are plotted with a dashed red line. 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.

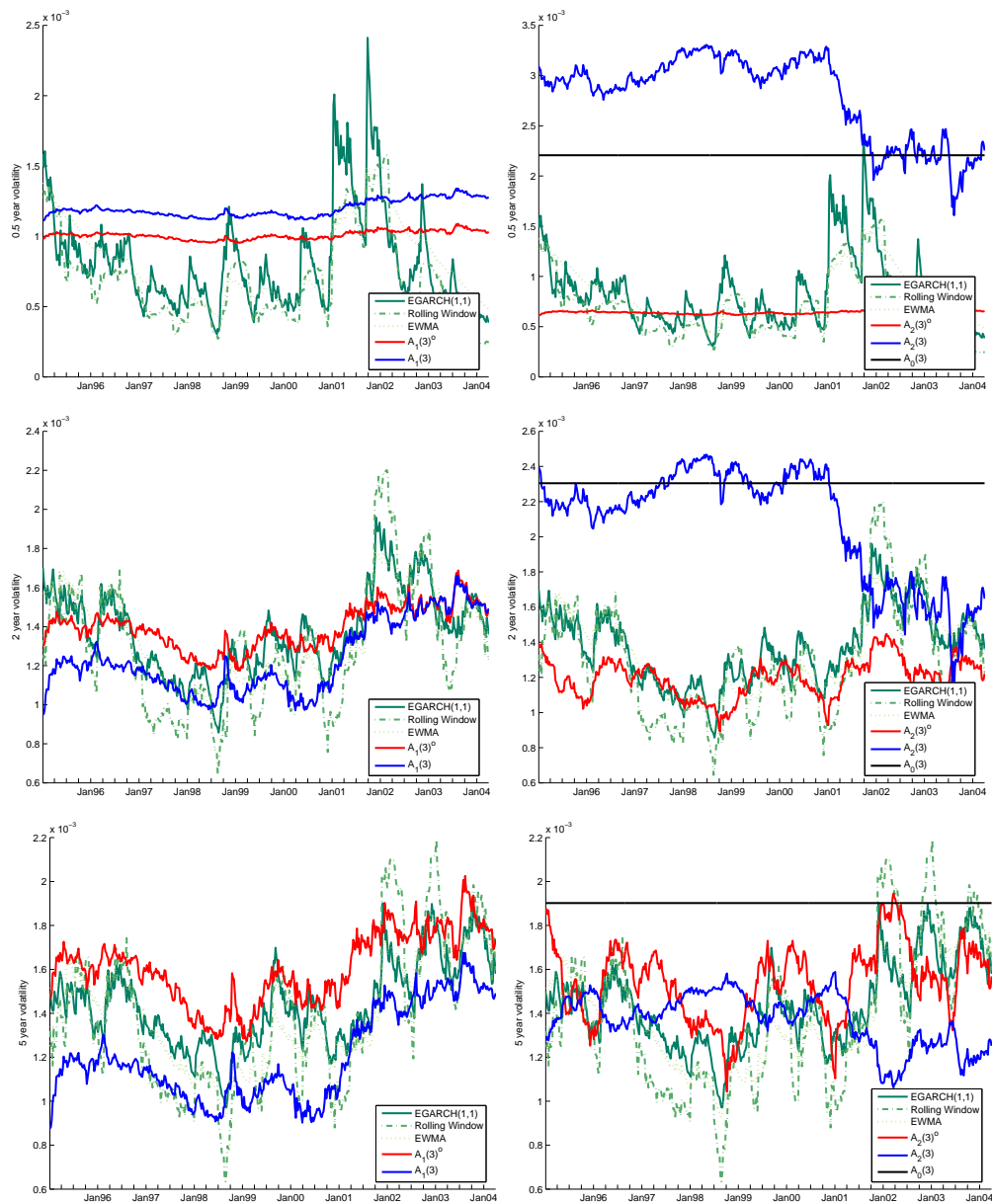


Figure 4: Realized Volatility

These figures plot model conditional volatility of zero coupon rates against various estimates of conditional volatility using historical data. For estimates of conditional volatility based on historical data we use a 26 week rolling window, an exponential weighted moving average (EWMA) with a 26-week half-life, and estimate an EGARCH(1,1) for each maturity. 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.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$	CP_5	CP_{10}	$CP_{5,10}$
2 Yr	-50.7	10.1	5.6	-2.3	9.8	22.2	-95.6	21.5
3 Yr	-40.5	22.7	6.7	5.2	8.1	27.1	-85.9	24.8
4 Yr	-39.2	24.1	10.7	11.7	2.8	31.8	-76.2	28.2
5 Yr	-40.8	25.9	13.8	16.7	-0.5	36.1	-66.6	31.7
6 Yr	-38.8	25.0	14.8	19.7	-3.0	0.0	-64.0	34.9
7 Yr	-37.7	28.5	15.7	22.7	-4.7	0.0	-64.0	33.8
8 Yr	-40.9	32.8	16.1	24.2	-5.2	0.0	-65.4	30.0
9 Yr	-42.7	34.1	15.8	25.1	-5.9	0.0	-68.9	27.9
10 Yr	-44.0	35.4	15.3	25.3	-6.4	0.0	-71.4	24.8

Table 8: Out-of-Sample Predictability of Excess Returns (R^2 's in %)

This Table presents R^2 's obtained from projections of weekly realized zero coupon returns, for different maturities, on model out-of-sample 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. $CP_{5,10}$ use 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 caps were measured with error.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$	CP_5
2 Year	48	25	10	17	18	37
3 Year	79	46	17	30	23	80
4 Year	107	69	23	42	24	119
5 Year	136	92	27	51	24	158
6 Year	165	114	30	58	23	
7 Year	195	134	32	64	22	
8 Year	224	153	34	68	21	
9 Year	253	170	36	72	21	
10 Year	282	185	38	76	21	

Table 9: Time Variation in Expected Returns.

This table contains the 1-week variance of 1-year expected excess return (expressed in basis points). 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. 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.

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
First Eigenvalue	1.18	1.53	2.29	1.62	1.81
Second Eigenvalue	0.34	0.57	0.17	0.17	0.64
Third Eigenvalue	0.03	0.57	0.00	0.13	0.04

Table 10: Eigenvalues of $\mathcal{K}_1^{\mathbb{P}}$ Matrix

	$A_0(3)$	$A_1(3)^o$	$A_1(3)$	$A_2(3)^o$	$A_2(3)$
First Eigenvalue	1.08	1.71	2.30	2.62	2.38
Second Eigenvalue	0.75	0.54	0.53	0.42	0.57
Third Eigenvalue	0.02	0.12	0.05	0.13	0.05

Table 11: Eigenvalues of $\mathcal{K}_1^{\mathbb{Q}}$ Matrix

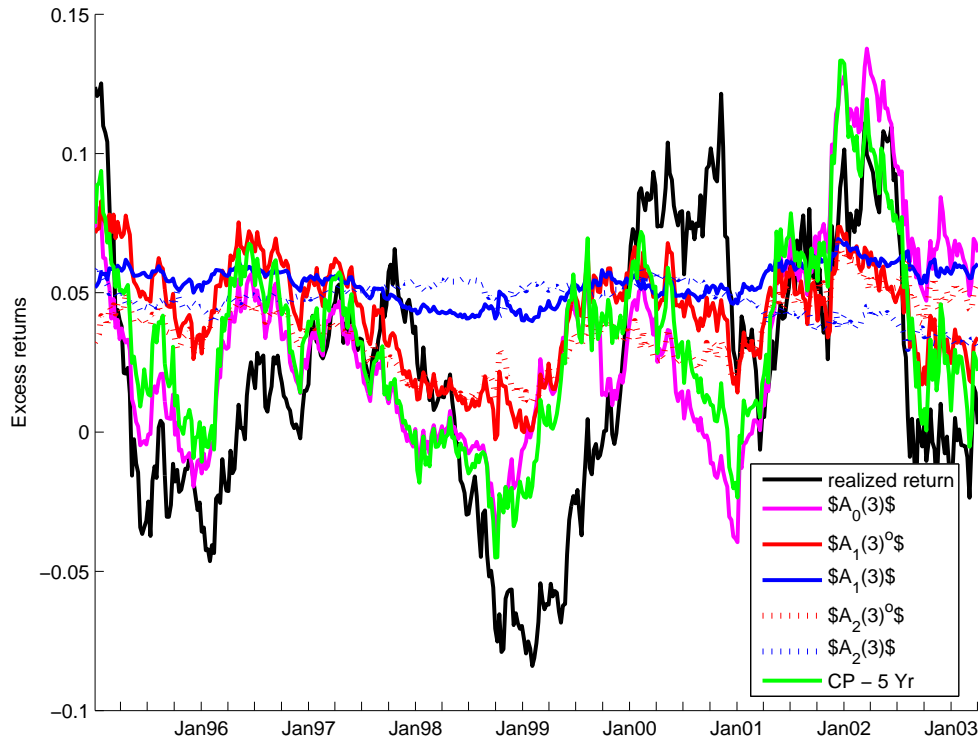


Figure 5: Excess Returns

This figure plots weekly realized excess returns, and model implied expected excess returns for a 5 year zero coupon bond. Realized excess returns are plotted with a solid black line. Predicted excess returns from the $A_0(3)$ model are plotted with a solid pink line. Predicted excess returns from the $A_1(3)$ model are plotted with a solid blue line and those from the $A_1(3)^o$ model are plotted with a solid red line. Predicted excess returns from the $A_2(3)$ model are plotted with a dashed blue line and those from the $A_2(3)^o$ model are plotted with a dashed red line. The prediction of excess returns from a regression of excess returns on 1-year zero rates and 1-year forward rates at 1-, 2-, 3-, and 4-years is labelled CP_5 and is plotted with a solid green line.