Current draft: June 8, 2017
First draft: August 8, 2005

Forecasting the Yield Curve Prior to the Global Financial Crisis: An Assessment of the Cox, Ingersoll, and Ross Model

Christopher G. Lamoureux
and
Kenneth D. Roskelley[1]

We find that the traditional expectations hypothesis holds for US Treasury bills during the period June 1989 - June 2007. Encompassing regressions of Treasury yield forecasts soundly reject the Cox, Ingersoll, and Ross model as the data generating process during this period. The model is of virtually no use as a forecasting tool: It produces forecast errors larger than the naive Martingale model. Furthermore these forecasts fail to provide incremental information about future yield changes and spreads. Our out-of-sample loss functions maintain the model's stochastic singularity, so the model's poor performance is not due to an unfortunate choice of an auxiliary error model. Instead rate dynamics are more complex than the model can accommodate.

JEL Classification: C90; D40; G10

[1]Department of Finance, The University of Arizona, Eller College of Business Administration, Tucson, 85721, 520–621–7488, lamoureu@lamfin.eller.arizona.edu; and Department of Finance and Economics, Mississippi State University, College of Business, Starkville, 39762, 662–325–1979, kdr149@msstate.edu, respectively. We retain full culpability and acknowledge intellectual debts to Jean Helwege, Kei Hirano, Jay Huang, George Jiang, Yuriy Kitsul, Doug Witte and participants at the 2013 Midwest Finance Association Meetings. This paper was previously titled, "Optimal Interest Rate Forecasts in the Greenspan Era." The current version of this paper may be downloaded from: *http://finance.eller.arizona.edu/lam/rsch.html*.

1. Introduction

Finance theory generally posits that interest rates are predictable. In this way they are very different from stock returns. For example the traditional expectations hypothesis models the six-month yield as the average of the three-month spot rate and the three-month rate in three months. Even when a liquidity premium drives a wedge between the two sides of this equation, the future three-month spot rate should be predictable, conditional on the three-month and six-month spot rates. In light of this many early and influential studies of term structure models evaluated models' predictive content. Fama (1976), Fama (1984), and Campbell and Shiller (1991) are classic examples. Dai and Singleton (2002) summarize these studies noting that a regression of the realized rate change on the change predicted under the expectations hypothesis at the three month horizon generates a coefficient of -0.428, with a standard error of 0.48. Under the expectations hypothesis this coefficient would be 1.

Considering the importance of interest rate predictability, both in practice, as well as for testing and evaluating model efficacy, we use forecast accuracy as our statistical loss function. We consider market yields on zero-coupon US Treasury securities from the period June 1989 through June 2007. This period precedes the global financial crisis, which is important since the behavior of interest rates in the period starting in late 2007 through 2016 is quite different from what it is in non-crisis periods. Our choice of time frame (a single operating policy regime at the Fed) and use of rolling regressions is motivated by past analysis of interest rate behavior. For example, Duffee (2006) finds evidence of a structural break in inflation dynamics and the term structure between 1952 and 1994, and Piazzesi (2005) stresses the importance of Fed policy in modeling the yield curve. Bansal and Zhou (2002) show that a two-factor CIR model with regime switches could cause the documented violations of the expectations hypothesis, over the period 1964 - 1995.

We find that during the 13 years prior to the global financial crisis, yield dynamics at the short end of the US Treasury yield curve conform to the traditional expectations hypothesis. This result is consistent with Longstaff (2000) who finds that general collateral repo rates conform to the pure expectations hypothesis, at up to the three month term, during the period May 1991 through October 1999. Longstaff attributes the sharp contrast with earlier rejections of the expectations hypothesis to his use of repo rates instead of yields on US Treasury securities, since the latter are influenced by liquidity and other security-specific features. Similarly, Downing and Oliner (2007) find that yields on commercial paper conform to the traditional expectations hypothesis in the period January 1998 to August 2003. Consistent with our results, Downing and Oliner suggest that a change in Federal Reserve operating policy in 1994–toward increased transparency–may be a reason for the dramatically improved performance of the expectations hypothesis. Poole and Rasche (2000, 2003), also suggest that our sample period covers a transparent Federal Reserve policy regime that has made it easier for the market to anticipate policy changes.

Modern affine term structure models (ATSMs) also generate predictable interest rates. These models start with a diffusion process for the state variable(s) that require mean reversion to prevent infinite interest rates. An important feature of affine term structure models is that they can produce forecasts of the entire yield curve. This allows us to consider forecasts not only of future rate changes but also various term spreads. Duffee (2002) suggests that the documented poor empirical fit of affine term structure models is due to their failure to reproduce the *failure* of the expectations hypothesis in the data. The fact that the expectations hypothesis holds in our data means that we can deepen our understanding of why ATSMs fare poorly empirically. The model forecast of future yields and spreads relies only on historical yield dynamics. Financial markets incorporate Federal Reserve policy guidelines and stated policies into interest rates. The information used by investors in determining the yield curve is therefore much richer than historical yield curves.

To support this contention, we design encompassing tests of the Cox, Ingersoll, and Ross (CIR) model, Aït-Sahalia's (1996) and Stanton's (1997) nonparametric models of interest rate dynamics, as well as the forward parity condition. The CIR model is attractive as it is a member of the class of ATSMs that generates heteroskedastic rate dynamics and allows the price of risk to vary with the state. We find that the models fare very poorly. Furthermore, we use the time-series tests of Giacommini and White (2006) to evaluate the out-of-sample predictive ability of the arbitrage model relative to other models using the root-mean-square error (RMSE) criterion. We show that CIR's RMSE is, in general, unconditionally and conditionally statistically significantly higher than the martingale's. Finally we ask whether there are market conditions under which the model performs relatively better as part of a forecast method. Here we find that there are no conditions in which CIR beats the martingale.

Our out-of-sample test design is appealing econometrically as we can test the no-arbitrage ATSM without adding an auxiliary model of errors. The addition of an arbitrary error model has important implications for empirical design, and has been analyzed by Renault (1997), Jacquier and Jarrow (2000), and Pastorello, Patilea, and Renault (2003). Moreover, Sims (2003) and Johannes and Polson (2003) warn of the econometric and decision-theoretic difficulties created by adding an auxiliary error model, in light of the high dimensionality of the error space and the fact that the underlying arbitrage pricing theory provides no guidance on constructing these errors.

Even if we ignore this joint hypothesis problem, Christoffersen and Jacobs (2004) caution that adding an error model gives rise to the need to maintain consistency between the role that errors play in estimation and inference when testing arbitrage-free models. They show that out-of-sample tests of option pricing models may perform poorly when the auxiliary model and estimator used at the estimation stage are not congruent with the quantity being forecast in the inference stage. So, while it is salutary to avoid distorting parameter estimates and implied state variables with an error model at the estimation stage, introducing an error model at the inference stage gives rise to this asymmetric loss function problem. Piazzesi (2010, esp. p.726) makes clear that this criticism applies to many classical studies.

Table 1 provides a synopsis of various approaches to estimate and test affine models. As is evident in the table, most of the empirical work on ATSMs has added an auxiliary error model. These error models may be justified on the basis that the model is only an approximation of the truth, or an acknowledgment that most empirical studies interpolate zero-coupon yields from coupon bonds. Nevertheless, these empirical tests are joint tests of the ATSM and the error model.[1] Like our methodology, Hong and Li (2005) do not add an error model at either the estimation or inference step. Their tests show that the inverse transition density of the implied state variables or yields have significantly different shapes and properties from what the model's transition density implies. We go one step further and show, using observed zero-coupon yields, that the model is not only a poor approximation for the underlying data generating process, but the model is also an uninformative forecasting method and an even worse approximation for yield dynamics than the Martingale model.

The remainder of the paper is organized as follows. The next section presents our data and the derivation of the alternative forecasts used in evaluating the ATSMs. Section 3 contains our results as we move sequentially through the four stages of inference. Section 4 concludes

[1] Research on term structure models has continued apace since the global financial crisis. Much of this recent research deals with the accommodative central bank policies in the aftermath of the crisis, see for example Swanson and Williams (2014) and Monfort, Pegoraro, Renne, and Roussellet (2017). This literature establishes that forecasting interest rates in the 10-year period following the end of our sample is a very different problem from this paper's focus, which is the behavior of US interest rates in non-crisis periods.

the paper.

2. Empirical Analysis

2.1 Data

We hand collect zero-coupon yields from the *Wall Street Journal* and Bloomberg. We record the closing ask yields-to-maturity on 3-month and 6-month Treasury bills, and approximately 5-, 15-, and 25-year principal STRIPS at a weekly frequency, on Wednesdays over the 942 week period starting on June 14, 1989 and ending on June 28, 2007. For those instances when Christmas and the following New Year fall on Wednesdays we use data from the preceding Tuesday. The US Treasury auctions 3- and 6-month bills every week, so the terms on these two yields are constant throughout our sample. Our 5-year STRIPS are derived from 5-year notes, which tend to be auctioned on a monthly basis, so the terms on these yields vary slightly throughout the sample. The 25- and 15-year STRIPS are derived from 30-year Treasury bonds whose availability is more restricted. This is especially true for the 15-year yield in the first half of our sample. There are two dates when we reset the term on the 15-year yield by more than one year: February 28, 1996 and December 3, 1997. In light of this, all *forecasts* that involve the 15-year yield exclude these two dates. Principal components analysis on these five yields show that the first eigenvalue accounts for 82.3% of the total variation, and the second accounts for 17%; the first two eigenvalues account for 99.3% of the total variation across the five yields.[2]

2.2 *Variables and Models Used in Forecasts*

2.2.1 CIR

Appendix A provides a summary of the CIR model. For the purpose of estimating a K-factor CIR model we use the most recent 250 (weekly) observations on the yields on K zero-coupon securities. As noted by Pearson and Sun (1994), this design implies a deterministic mapping from the yields to the factors. So by including the Jacobian term in the likelihood we obtain maximum likelihood estimates of the parameters governing factor dynamics and the risk premia associated with the factors.

[2] Rudebusch and Tao (2008) also find that two factors are sufficient to model monthly US Treasury yields during the 1988 - 2000 period. They suggest that this is because the sample period does not include "the period of heightened interest rate volatility during the late 1970s and early 1980s;" (p. 909).

Armed with the parameters and implied factor(s) on date t, we construct the $\tau-$step ahead forecast of factor z_j $(j = 1, \cdots, K)$ by:

$$E(z_{j,t+\tau}) = z_{j,t}e^{-\kappa_j \cdot \tau} + \theta_j(1 - e^{-\kappa_j \cdot \tau}) \tag{1}$$

Where, κ_j and θ_j are parameters that govern the dynamics of factor j (in the physical measure, as shown in Appendix A). Under the model the factors evolve independently and, as shown in Appendix A, yields at time t are linear functions of the factors. Note that the forecasted yields are priced without error on every date in the estimation period. Under the model's assumptions we construct the expected future yield(s) conditional on the model and rolling sample using (A.4).[3]

The likelihood function is neither smooth nor unimodal, making global optimization problematic. We use an optimization algorithm that starts with simulated annealing (following Goffe, Ferrier, and Rogers 1994). After the simulated annealing procedure converges, we use a gradient-based optimization in the neighborhood of this optimum. We repeat this procedure using three different starting conditions and parameter bounds, and take the maximum.

2.2.2 Nonparametric Forecasts

At the 1-week horizon for the 3-month yield we also include forecasts from the nonparametric (short-rate) models of Aït-Sahalia (1996) and Stanton (1997). Stanton (1997) uses the Nadaraya-Watson kernel estimator as follows:

$$\hat{\mu}(\xi_i) = \frac{1}{\Delta t} \frac{\sum_{t=1}^{T-1} (x_{t+1} - x_t)\phi\left(\frac{\xi_i - x_t}{h}\right)}{\sum_{t=1}^{T-1} \phi\left(\frac{\xi_i - x_t}{h}\right)} \tag{2}$$

Here, $\mu(\xi_i)$ is the expected change in yield from t to t+1, evaluated at $\xi_i, i = 1, \ldots N$ which is an equally spaced grid over the support of yields over the past 250 weeks. The x_t are the observed yields (in levels) over the sample. We use $N = 18$ grid points. We follow Stanton (as described in Chapman and Pearson 2000, p.360) and set the bandwidth, $h = 4\hat{\sigma}T^{-\frac{1}{5}}$, where T is the length of the period used in estimating the model (250 weeks), and $\hat{\sigma}$ is the sample standard deviation of the changes in yields over this estimation period. $\phi(\cdot)$ is the standard unit normal probability density function. The 1-week ahead forecast is obtained by estimating the 18 grid points $\hat{\mu}(\xi_i)$, and then using linear interpolation to pin down the yield at date t

[3]There is no Jensen's Inequality problem since the yield is linear in the state variables.

within the range of ξ_1 to ξ_N, and adding the forecast change to this last yield in the estimation period.

Aït-Sahalia (1996) posits an alternative nonparametric diffusion estimator. Since our focus is only on forecasting the future rate, we use the μ function that he chose:

$$\mu(x_t) \quad = \quad \alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2 + \alpha_3 \frac{1}{x_t}, \tag{3}$$

where $\mu(x_t)$ is the expected change in yield from t to t+1, and x_t represents the level of the yield at time t.

Note that both Aït-Sahalia and Stanton's estimators resemble Dickey-Fuller regressions: We project changes in rates on the lagged level. Both nonparametric methods are single factor models of short rate dynamics that allow for non-linear mean reversion.

2.2.3 Litterman and Sheinkman State Variables

We include the traditional yield curve factors as conditioning state variables in our analysis. These are the level, slope and curvature of the yield curve. We measure the level as the 3-month yield; the slope as the spread between the 25-year yield and the 3-month yield; and curvature as the difference between the sum of the the 3-month and 25-year yields and twice the 5-year yield. These variables were identified as yield curve factors by Litterman and Sheinkman (1991). Diebold and Li (2006) show that these state variables have predictive content for yields at horizons of two to four quarters.

2.2.4 The Forward Rate

The six-month spot rate is the average of the three-month spot rate and the three-month forward rate. As noted in the introduction, the expectations hypothesis identifies the three-month forward rate with the market's expectation of the three-month spot rate three months hence. We follow Longstaff (2000) and make no distinction between the different forms of the expectations hypothesis.[4] We regress the actual change in the 3-month spot rate from day t to

[4]Campbell (1986) shows that the empirical differences between alternative versions of the expectations hypothesis are imperceptible at short maturities.

day t+90 on the 3-month forward rate at t less the 3-month spot rate at t, for the 693 weeks in our "out-of-sample" data, as in Campbell and Shiller (1991).[5] The regression coefficient on the independent variable in this regression is 0.85, with a (Newey-West) standard error of 0.22. Thus in our sample, we cannot reject the traditional expectations hypothesis. When we add the realized changes in the Federal Reserve's target federal funds rate to this regression, the coefficient on this variable is 0.84 with a standard error of 0.06, and the coefficient on the difference between the forward rate and spot rate falls to .08 with a standard error of .09. While not formal (since we cannot argue that future Fed policy changes are exogenous), this suggests that changes in the target federal funds rate are a significant driver of 3-month yields, and that forward rates are a noisy proxy for these changes.

3. Results

Figure 1 shows the behavior of the 3-month yield over our out-of-sample period, along with the predictions from the two-factor CIR model, the nonparametric models of Aït-Sahalia and Stanton, and the Federal Reserve's stated target federal funds rate. This figure shows that all models do a good job at forecasting 3-month interest rate *levels*, and these levels closely follow the target federal funds rate. To avoid the well-known problems associated with forecasting persistent series we focus on yield changes.

3.1 *Encompassing Regressions*

Inference at our first two stages relies on encompassing regressions (Fair and Shiller 1990). The 250-week rolling estimation period means that our out-of-sample analysis covers the 693 weeks from week 251 (March 31, 1994) through the end of the sample. We conduct all inference at three forecast horizons: 1-week, 4-week, and 13-week. Since the residuals in these regressions are neither normally distributed nor independent under the null hypothesis in either Stage 1 or Stage 2, we estimate the Newey-West (1987) heteroskedasticity and autocorrelation consistent (HAC) variance-covariance estimators. For all regressions we also report the bandwidth from Andrews' (1991) procedure, which is the lag length used in adjusting the autocorrelations in the standard errors.

[5] We measure the "3-month forward rate three months hence" on date t as follows: $F_t^{3m} = \frac{R_t^{6m} \cdot \tau_{6m} - R_t^{3m} \cdot \tau_{3m}}{\tau_{6m} - \tau_{3m}}$. (Where R_t^{6m} is the 6-month yield on date t, R_t^{3m} is the 3-month yield on date t, τ_{3m} and τ_{6m} are the terms, in years, of the 3- and 6-month yields, respectively.) Under the traditional pure expectations hypothesis, this corresponds to the expected value of the 3-month rate in three months.

3.1.1 Yield Changes

Encompassing regression results for changes in all individual yields are reported in Table II. The baseline encompassing regressions are:

$$
\begin{aligned}
R^i_{t+\tau} - R^i_t = & \alpha + \beta_{CIR}(\hat{R}^{i,CIR}_{t+\tau} - R^i_t) + \beta_m(\hat{R}^{3m,m}_{t+\tau} - R^i_t) \\
& + \beta_l l_t + \beta_s s_t + \beta_c c_t + \beta_f \hat{f}_t + \epsilon_t
\end{aligned} \tag{4}
$$

where: R^i_t is the spot rate at time t on a zero coupon bond that matures in $i-$periods. $\hat{R}^{i,CIR}_{t+\tau}$ is the CIR model's forecast conditional on information available at time t of the $i-$period spot rate at time t+τ, and $\hat{R}^{3m,m}_{t+\tau}$ is the similarly defined forecast from the nonparametric models of Aït-Sahalia (1996) and Stanton (1997) at the 1-week forecast horizon. Note that these latter two models are only used to forecast the 3-month yield. The additional (exogenous) state variables introduced in the preceding section, are: the level, l_t, slope, s_t, and curvature, c_t, of the yield curve on date t; and $\hat{f}_t = F^{3m}_t - R^{3m}_t$, the difference between the 3-month forward rate, three months hence, and the 3-month spot rate on date t.

Stage 1 and Stage 2 inference may be formalized as:

$$\text{Stage 1} : \mathrm{H_0} :\ \alpha = 0; \beta_{\mathrm{CIR}} = 1; \beta_{\mathrm{m}} = 0$$

$$\text{Stage 2} : \mathrm{H_0} :\ \beta_{\mathrm{CIR}} = 0;\ \mathrm{H_A} : \beta_{\mathrm{CIR}} > 0$$

To avoid introducing an error, our research design requires that the forecasted rate be used in estimating the model, so when estimating a two-factor model we must choose a (single) companion rate. In this case, each of the possible pairs provides a distinct forecast. We construct forecasts using all possible pairs. To preserve space, we report results using the companion yield that provides the highest t-statistic on the CIR model's slope coefficient at the one week ahead forecast horizon ($\tau = 1$) encompassing regression.[6] For formal inference

[6] If the ATSM were true, its stochastic singularity implies that all choices of companion rate would result in identical forecasts. We are willing to entertain the situation in which the model is rejected, to explore whether it can serve as tool to extract useful forecasting information from the data. In addition to stochastic singularity, there are identification issues that inform this procedure. For example, if we estimated the two-factor model using only the two T-Bill yields, the likelihood surface tends to be flatter than in cases where the model is estimated using yields on a bill and a STRIPS. Longer term rates are especially informative in terms of identifying the risk premia.

we use Monte Carlo to construct the density of the maximum of four $t-$statistics under the null hypothesis.[7]

The dependent variable in Table II, Panel A is the change in the 3-month yield. For ease of interpretation the $t-$statistics are for the Stage 2 null hypothesis that the coefficient equals zero. Nevertheless, it is obvious that we reject the Stage 1 hypothesis that the intercept is zero and the coefficient on CIR is one. The model's Stage 1 orthogonality restriction is also uniformly rejected as the coefficient on the alternative exogenous variables, the level and the difference between the forward rate and spot rate, are uniformly statistically significant. For this yield the only case where the CIR model contains incremental information in the encompassing regression (i.e., Stage 2 inference) is when the one-factor model is used to construct a forecast at the 1-week horizon. Overall, the method of Aït-Sahalia provides useful forecasting information at the 1-week horizon, although the sign is negative. Here we also see that the coefficient on the forward rate minus the spot rate at the 13-week horizon is not statistically different from unity–even in the presence of the other state variables and CIR forecast. This reinforces the empirical support for the traditional expectations hypothesis in our sample period.

The results for the 6-month yield, shown in Table II, Panel B, are very similar to those for the 3-month yield. The main exceptions are that the coefficient on the CIR model is never statistically significant, curvature is significant at the 1-week horizon, and the difference between the 3-month forward rate and the spot rate is only significant at the 13-week horizon. Turning to the three STRIPS yields in Table II, Panels C, D and E, we see that the only case where CIR is a significant predictor (with a positive coefficient) is for the two-factor model's forecast of the 5-year yield at the 1-week horizon (Panel C). Unlike the two shorter-term rates, the level and the forward minus spot are not significant (at the 5% level) in any of these encompassing regressions. The slope is never statistically significant. Indeed for the 15- and 25-year yields none of the exogenous regressors is significant at any horizon.

Figure 2 provides some insight to the models' forecast performance. The figure charts forecasts of the 3-month rate at the 13-week horizon. We see in Figure 1 that the 3-month yield is gradually and monotonically increasing until late 1995, reflecting the rising federal funds target rate. By contrast, in Figure 2 the CIR model predicts that the 3-month rate

[7] These distributions are constructed with 10 million draws. We construct the $p-$values for one-sided tests. The 5% critical value for the maximum of four t-statistics is 2.24 for a one-tailed test (and 2.50 for the corresponding two-tailed test). We also use these critical values for the 4- and 13-week forecast horizons. In all analysis, we use a one-tail test (that the coefficient is significantly positive) for CIR and two-tail tests for all other exogenous forecasts.

will fall over this same period, as the model anticipates rates will regress to the sample mean. As Clarida, Gali, and Gertler (2000) and others note, however, target rates exhibit inertia as monetary policy is implemented smoothly over time. Hence, as the Fed's target rate gradually falls in late 1994, CIR forecasts start to rise. In general, as the target federal funds rate changes, CIR predicts the short rates will move oposite the federal funds rate changes. CIR forecasts of the short rates fail to account for monetary policy inertia.

3.1.2 Changes in Term Spreads

Cheridito, Filipović and Kimmel (2007) and Duarte, Longstaff, and Yu (2007) suggest that ATSMs may do a better job at forecasting the dynamics of a cross-section of yields than a single yield. So in this sub-section we evaluate forecasts of the spreads between all of the yield pairs in our sample. Forecasting a function of two yields removes the question of which yield pair should be used to estimate the model; both yields are used to ensure that the model fits the data without error on the date the forecast is constructed. We report the following encompassing regressions in Table III:

$$\begin{aligned}(R^i_{t+\tau} - R^j_{t+\tau}) - (R^i_t - R^j_t) =& \alpha + \beta_{CIR}\left[(\hat{R}^{i,CIR}_{t+\tau} - \hat{R}^{j,CIR}_{t+\tau}) - (R^i_t - R^j_t)\right] \\ &+ \beta_l l_t + \beta_s s_t + \beta_c c_t + \beta_F \hat{f}_t + \epsilon_t \end{aligned} \tag{5}$$

where all variables are as defined above. Here the variable being forecast is the spread between two rates at a $\tau-$ step ahead horizon. Stage 1 inference soundly rejects the model as the only case (of the 30 considered) where the coefficient on the CIR forecast spread is statistically larger than zero, and not distinguishable from one, is the spread between the 25-year and 15-year yields at the 4-week horizon. But in this case, the coefficient on the curvature is also significant. As for Stage 2, the only other case where the coefficient on the CIR forecast is statistically larger than zero is for the spread between the 5-year and 6-month yields at the 1-week horizon, and this coefficient is significantly smaller than one. Also the coefficients on the slope and curvature are statistically significant in this regression.

Since we saw in Table II that the difference between the forward rate and the spot rate predicts the 3-month yield it is not surprising that this variable has significant predictive content when forecasting the spreads between longer yields and this short rate. In Panel D, for example, which considers the spread between the 25-year yield and the 3-month yield, if the forward rate is 100 basis points above the spot rate, then we expect that the spread between the 25-year and 3-month yields will decline by 102 basis points over the next 13 weeks. The convexity of the yield curve as measured by curvature also has predictive content for several of our pairwise yield spreads.

3.2 Conditional Predictive Ability of Forecasting Methods

In light of formal rejection at Stages 1 and 2, our interest shifts to whether the model has any information as a forecasting tool. Stage 2 encompassing regressions are not conclusive on this point for several reasons. First, since the forecast weights are estimated, the Stage 2 analysis of the model's *incremental* information is evaluated ex post. Furthermore these weights are affected by the correlations between competing forecasts. To be truly out-of-sample, as well as to allow consideration of the *total* information content of the model, we test the arbitrage model's predictive ability by comparing the relative sizes of forecast errors from the model with those from the martingale model. Meese and Rogoff (1983) suggest that an informal comparison of the RMSE from structural models to that from a martingale model is a useful assessment of the model's economic viability. A comparison to the martingale model is attractive in that an informative model should be able to outperform a forecast that the future rate will equal its current level. While Cheridito, Filipović, and Kimmel (2007) and Duffee (2002) report RMSEs of the martingale model and various ATSMs, neither study provides a formal, statistical comparison of the competing models' RMSEs. Diebold and Mariano (1995) and West (1996) develop a statistical framework that affords a formal comparison the RMSEs of two models. Clark and West (2006) show that these tests may be biased when the martingale model is included in the analysis. Because we have already formally rejected the CIR model, we use the approach proposed by Giacomini and White (2006) which, in addition to circumventing this problem, allows for model mis-specification and is robust to unmodeled heterogeneity.

The Giacomini and White test formally compares the predictive ability of alternative forecasting methods. Our Stage 3 inference evaluates whether the arbitrage model can produce a forecast that is at least as good as the naive martingale model. Whether we reject at Stage 3 or not, it is conceivable that alternative models produce the best available forecast under certain market conditions. The Giacomini and White test allows us to analyze whether CIR as a forecasting method produces informative forecasts under specific market conditions. We characterize those conditions in our Stage 4 analysis.

Define $\Delta L_{t+\tau}$ as the difference in the squared error of the two models' forecasts of the yield or spread at time t+τ, conditional on the information set at time t, $\mathcal{F}_t$. The null hypothesis is that, conditional on $\mathcal{F}_t$ the difference between the predictability of the two models is zero: $E\left[\Delta L_{t+\tau}|\mathcal{F}_t\right] = 0$. As Giacomini and White (2006) point out, the null hypothesis that $\Delta L_{t+\tau}$ is a martingale difference sequence allows us to write an orthogonality restriction that $E\left[h_t \Delta L_{t+\tau}\right] = 0$ for all $\mathcal{F}_t$ measurable functions h_t. Using a $q-$dimensional vector, h_t, Giacomini and White's Wald statistic has the form

$$T^h_{m,n} = n\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right)' \widehat{\Omega}^{-1}\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right) \tag{6}$$

where n is the number of (overlapping) observations, m is the size of the rolling estimation window, and $\widehat{\Omega}$ is the estimated covariance matrix of the sample conditional moment restriction, $\frac{1}{n}\sum\limits_{t=m}^{T-1} h_t \Delta L_{t+\tau} = 0$. The test statistic is asymptotically distributed as chi-square with q degrees of freedom. As with the encompassing regressions, we construct $\widehat{\Omega}$ using Andrew's (1991) bandwidth procedure to identify the lag length used to calculate the Newey-West (1987) covariance estimator.

The performance of this χ^2 test statistic depends on the choice of h_t. We specify $h_t^{'} = \left\{1, l_t, s_t, c_t, \hat{f}_t\right\}$, and report the statistic for both the unconditional, ($h_t^{'} = \{1\}$) and conditional tests.

$$\mathrm{Stage\,3:\ H_0:\ E[\Delta L_{t+\tau}\mid \mathcal{F}_t] = 0}$$

$$\mathrm{Stage\,4:\ Estimate\,E[h_t - \bar{h}\mid \Delta L_{t+\tau} < 0]}$$

Table IV contains all of the RMSEs for forecasts of the change in each interest rate from all of the models discussed above and the martingale. For each yield, at each horizon, the forecast with the lowest RMSE is in bold face. We test the null hypothesis that the model produces a forecast that is no more accurate than the martingale using the Giacomini and White conditional χ^2 test statistic. We report the unconditional test statistic in parentheses and the conditional test statistic in brackets below the RMSE.

The ATSM fares poorly in our Stage 3 tests. A CIR model produces a lower RMSE than the martingale in only one of the 15 yield/horizon cases; the 13-week horizon forecast of the 25-year yield (Panel E). For this one case, estimating CIR with the 25-year yield and either the 6-month or 5-year yield results in a lower RMSE than the martingale, but neither of these differences is statistically significant. By contrast, at both shorter horizons the martingale RMSE is statistically smaller than all CIR models. In fact, the only other cases where the CIR model's RMSE is not significantly larger than the martingale's are at the 1-week horizon for the 5- (Panel C) and 15-year (Panel D) yields.

In Panel A, the 3-month forward rate at the 13-week forecast horizon also has a significantly smaller RMSE than any CIR model, though at the 1-week horizon the CIR models produce forecasts with significantly lower RMSEs than either nonparametric forecast. For these two short rates (Panel A and Panel B) the one-factor CIR model forecasts have smaller RMSEs than those from all two-factor CIR models at all three horizons. Amongst the three longer-term yields the one-factor model has a lower RMSE than all two-factor models only for the 1- and 4-week horizon forecasts of the 25-year yield.

Turning again to the models' cross-sectional restrictions, in Table V we consider statistical comparisons of the RMSEs of all pairwise spread forecasts. The martingale produces a significantly smaller RMSE in 26 of the 30 cases, including all three horizons for the following yield pairs: 5-year and 3-month, 15-year and 3-month, 5-year and 6-month, 15-year and 6-month, and 25-year and 6-month. The only case where the CIR model produces a significantly lower RMSE than the martingale is the 4-week horizon forecast of the spread between the 25-year and 15-year yields.

Interestingly there is only one case in Table V, and ten cases in Table IV where qualitative inference is affected by the choice of h_t (assuming a critical value of 5%). Generally, adding the four state variables to h_t reduces the test's power–so we reject the null unconditionally, but fail to reject conditionally. The exception–where the state variables increase the test's power–is for the comparison between the squared forecast errors from the martingale and difference between the forward rate and spot rate for the 3-month yield forecast at the 13-week horizon.

3.3 Conditional Predictive Ability: Decision Rule Assessment

Giacomini and White (2006, p.1569) suggest using the state variables in a decision rule assessment to enhance the information contained in the test. Since Stage 3 is like Stages 1 and 2 in soundly rejecting the ATSM, we now ask whether there are any conditions under which the arbitrage model does better than the martingale. We follow Giacomini and White (2006) and project ΔL_t onto h_t to ascertain the conditions under which one model is preferred to another. In general, we find that the ATSM does relatively better in periods when the level variable (i.e., the short rate) and slope are relatively high and when the curvature and spread between the forward and spot rates are relatively low. However, the overarching conclusion is that the martingale beats the CIR model uniformly.

Table VI and Figure 3 provide more information on assessment of the model forecast errors. In Table VI we report the effect of a one standard deviation increase in each of the yield curve state variables for those cases where the coefficient on the state variable is statistically significant in the regression of ΔL_t on h_t. A glance at this table reveals that for all three long rates and for most of the yield spreads, there is no statistically significant relationship between the relative RMSEs (of CIR and the martingale) and the state variables. The situation is different for the two shorter-term yields. Consider the 13-week ahead forecast of the 6-month yield. In this case the RMSE of the one-factor CIR model is 4.237 basis points higher than that from the martingale. A one standard deviation decrease in the spread between the spot and forward rate would result in the CIR model's RMSE being 2.076 basis points lower than the martingale. The only other cases in Table VI where the CIR RMSE falls below the martingale RMSE are for the 13-week ahead forecast of the 3-month yield, where a one standard deviation increase in this yield or a one standard deviation decrease in the forward rate minus this yield bring the CIR RMSE below the martingale's. The only yield spread where the relationship between the squared errors is state-dependent is that between the 15- and 5-year yields. The martingale RMSE is lower than that from the two-factor CIR by 0.914 basis points. When the 90-day yield is one standard deviation above its mean, this difference falls to 0.178 basis points (but the martingale is still better).

Figure 3 provides a graphical representation of our Stage 4 assessment for the mean square error from the indicated CIR model minus the mean square error from the martingale forecasts for: the 3-month yield at the 1-week forecast horizon plotted against the slope (Panel A), the 3-month yield at the 4-week horizon plotted against the forward rate minus the level (Panel B), the 6-month yield at the 13-week horizon plotted against curvature (Panel C), and the spread between the 15-year and 5-year yields at the 13-week horizon plotted against the level (Panel D). In Panel A we see that the martingale and CIR perform similarly when the yield curve's slope exceeds 4% and when the yield curve is downward sloping.

In Figure 3, Panel B, we see that CIR forecasts are much worse than those from the martingale when the spread between the 3-month forward exceeds the 3-month yield by more than 75 basis points. We can see from Figures 1 and 2 that this occurs most commonly in the first year of our out-of-sample period, when both the 3-month yield and the target federal funds rate are generally rising, whereas the CIR model is forecasting a drop in the short rate. In Panel C we see that the CIR forecast of the 6-month yield in 13 weeks is better than the martingale when curvature is highest. CIR does worse when the yield curve is linear (i.e., curvature is zero), and when there is negative convexity. In Panel D we see that both the CIR and martingale forecasts of the spread between the 15-year and 5-year yields exhibit a high variance when the 3-month yield is less than 2%. In Figure 1, we see that this is the period following the September 11, 2001 disruptions through early 2005. By contrast, CIR tends to outperform the martingale forecast of this portion of the yield curve's slope when the 3-month yield exceeds 5%.

4. Conclusion

We design a sequence of tests that do not entail adding an auxiliary error model to an arbitrage-free affine term structure model, and which allow us to test the model as the data generating process as well as a forecasting tool. This allows us to reject the notion that the model's poor empirical performance is the result of an unfortunate choice of error model. We soundly reject the model as the data generating process, and find the model provides little to no incremental information about future yield changes or yield spreads. Similarly, as a forecasting tool the model fares poorly, producing root mean square errors that are statistically larger than those of the Martingale model.

Interest rate behavior in our sample is different from that in earlier periods. Our sample is restricted to the post-Greenspan era, where Longstaff (2000) and Downing and Oliner (2007) provide evidence that the expectations hypothesis holds, in part due to transparent monetary policy. Clarida, Gali, and Gertler (2000) and others document that monetary policy is implemented gradually, through successive increases or cuts to the target policy rate. Affine models fail to account for these complex dynamics, and instead forecast reversion to the historical mean as the policy rate changes. Duffee (2002) suggests that the poor empirical fit of affine term structure models is due to their failure to reproduce the *failure* of the expectations hypothesis in the data. As we also find support for the expectations hypothesis, the failure of the CIR model in this study complements Duffee's result. Here the CIR model fails at the short-end of the yield curve because it cannot reproduce the success of the traditional expectations hypothesis in our data.

Appendix A. The CIR Model

Cox, Ingersoll, and Ross (1985) model an equilibrium, no-arbitrage economy, with a representative agent. From a theoretical point of view, the model links the data on yields to one or more latent independent factors. The model posits that the time series evolution of this latent factor is a mean-reverting, square-root process:

$$dz_j = \kappa_j(\theta_j - z_j)dt + \sigma_j\sqrt{z_j}d\omega_j, \tag{A.1}$$

where:
$j = 1, \cdots, K$ (the number of factors), and
ω_j is a Wiener process.

Feller (1951) shows that the transition density for any z_j at time $t+\tau$ conditional on its realization at time t is given by: (suppressing the j subscript)

$$f(z_{t+\tau}|z_t) = c\, e^{-u-\nu}\left(\frac{\nu}{u}\right)^{q/2} I_q(2(u\nu)^{1/2}), \tag{A.2}$$

where:
$c = \frac{2\kappa}{\sigma^2(1-e^{-\kappa\cdot\tau})}$
$u = cz_te^{-\kappa\cdot\tau}$
$\nu = cz_{t+\tau}$
$q = \frac{2\kappa\theta}{\sigma^2} - 1$

I_q is a modified Bessel function of the first kind of order q.

Bond prices depend on the current value of the state variable, as well as its expected evolution, along with a risk premium, λ. Specifically, the price of a τ-year bond, at time t is:

$$P_{t,t+\tau} = \prod_{j=1}^{K} \Lambda_{j,t,\tau}\; e^{-\beta_{j,t,\tau}\cdot z_{j,t}}, \tag{A.3}$$

where:
$\Lambda_{j,t,\tau} = \left[\frac{2\gamma_j e^{[(\kappa_j+\lambda_j+\gamma_j)\tau]/2}}{(\kappa_j+\lambda_j+\gamma_j)(e^{\tau\gamma_j}-1)+2\gamma_j} \right]^{2\kappa_j\theta_j/\sigma_j^2}$
$\beta_{j,t,\tau} = \frac{2(e^{\tau\gamma_j}-1)}{(\kappa_j+\lambda_j+\gamma_j)(e^{\tau\gamma_j}-1)+2\gamma_j}$
$\gamma_j = ((\kappa_j + \lambda_j)^2 + 2\sigma_j^2)^{1/2}$.
For zero coupon bonds, the continuously compounded yield to maturity is:

$$R_{t,t+\tau} = \frac{\sum_{j=1}^{K} (\beta_{j,t,\tau} \cdot z_j - \log \Lambda_{j,t,\tau})}{\tau}. \tag{A.4}$$

References

Aït-Sahalia, Yacine, 1980, Testing continuous-time models of the spot interest rate, *Review of Financial Studies* **9**, 385–426.

Andrews, Donald, 2002, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817–858.

Bams, Dennis and Peter C. Schotman, 2003, Direct estimation of the risk-neutral factor dynamics of Gaussian term structure models, *Journal of Econometrics* **117**, 179–206.

Bansal, Ravi and Hao Zhou, 2002, Term Structure of interest rates with regime shifts, *Journal of Finance* **57**, 1997–2043.

Brown, Stephen J. and Philip H. Dybvig, 1986, The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates, *Journal of Finance* **41**, 617–630.

Campbell, John, 1986, A defense of traditional hypotheses about the term structure of interest rates, *Journal of Finance* **41**, 183–193.

Campbell, John and Robert Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* **58**, 495–514.

Chapman, David A. and Neil D. Pearson, 2000, Is the short rate drift actually nonlinear? *Journal of Finance* **55**, 355–388.

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

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

Christoffersen, Peter and Kris Jacobs, 2004, The importance of the loss function in option valuation, *Journal of Financial Economics* **72**, 291–318.

Clark, Todd E. and Kenneth D. West, 2006, Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis, *Journal of Econometrics* **135**, 155–186.

Clements, Michael P. and David F. Hendry, 1993, On the limitations of comparing mean square forecast errors, *Journal of Forecasting* **12**, 617–637.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, A theory of the term structure of interest rates, *Econometrica*, **53**, 385–407.

Dai, Qiang and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* **55**, 1943–1978.

Dai, Qiang and Kenneth J. Singleton, 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* **63**, 415–441.

DeJong, Frank, 2000, Time series and cross-section information in affine term structure models, *Journal of Business and Economic Statistics* **18**, 300–314.

Diebold, Francis X. and Canlin Li, 2006, Forecasting the term structure of government bond yields, *Journal of Econometrics* **130**, 337–364.

Diebold, Francis X. and Roberto S. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* **13**, 253–263.

Downing, Chris and Stephen Oliner, 2007, The term structure of commercial paper rates, *Journal of Financial Economics* **83**, 59–86.

Duarte, Jefferson, Francis A. Longstaff, and Fan Yu, 2007, Risk and return in fixed income arbitrage: Nickels in front of a steamroller? *Review of Financial Studies* **20**, 769–811.

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

Duffee, Gregory R. and Richard H. Stanton, 2012, Estimation of dynamic term structure models, *Quarterly Journal of Finance* **2**.

Duffie, Darrell and Ken Singleton, 1997, An econometric model of the term structure of interest-rate swap yields, *Journal of Finance* **52**, 1287–1321.

Egorov, Alexei, Yongmiao Hong, and Haitao Li, 2006, Validating forecasts of joint probability density of bond yields: Can affine models beat random walk? *Journal of Econometrics* **135**, 255–284.

Fair, Ray and Robert Shiller, 1990, Comparing information in forecasts from econometric models, *American Economic Review* **80**, 375–389.

Fama, Eugene F., 1976, Forward rates as predictors of future spot rates, *Journal of Financial Economics* **3**, 361–377.

Fama, Eugene F., 1984, The information in the term structure, *Journal of Financial Economics* **13**, 509–528.

Fama, Eugene F., and Michael R. Gibbons, 1984, A comparison of inflation forecasts, *Journal of Monetary Economics* **13**, 327-348.

Feller, William, 1951, Two singular diffusion problems, *Annals of Mathematics* **54**, 173–182.

Giacomini, Raffaella and Halbert White, 2006, Tests of conditional predictive ability, *Econometrica* **74**, 1545–1578.

Gibbons, Michael R. and Krishna Ramaswamy. 1993, A test of the Cox, Ingersoll, and Ross model of the term structure, *Review of Financial Studies* **6**, 619–658.

Goffe, William L., Gary D. Ferrier, and John Rogers, 1994, Global optimization of statistical functions with simulated annealing, *Journal of Econometrics* **60**, 65–99.

Hong, Yongmiao and Haitao Li, 2005, Nonparametric specification testing for continuous-time models with applications to term structure of interest rates, *Review of Financial Studies* **18**, 37-84 .

Jacquier, Eric and Robert Jarrow, 2000, Bayesian analysis of contingent claim model error, *Journal of Econometrics* **94**, 145–180.

Johannes, Michael and Nicholas Polson, 2003, Comment (On Iterative and recursive estimation in structural nonadaptive models), *Journal of Business and Economic Statistics* **21**, 500-503.

Lamoureux, Christopher G. and H. Douglas Witte, 2002, Empirical analysis of the yield curve: The information in the data viewed through the window of Cox, Ingersoll, and Ross, *Journal of Finance* **57**, 1479–1520.

Litterman, Robert and José Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* **1**, 54–61.

Longstaff, Francis A., 2000, The term structure of very short-term rates: New evidence for the expectations hypothesis, *Journal of Financial Economics* **58**, 397–415.

Meese, Richard and Kenneth Rogoff, 1983, Empirical exchange rate models of the seventies, *Journal of International Economics* **14**, 3–24.

Monfort, Alain, Fulvio Pegoraro, Jean-Paul Renne, and Guillaume Roussellet, 2017, Staying at zero with affine processes: An application to term structure modelling, *Journal of Econometrics*, forthcoming.

Newey, Whitney and Ken West, 1987, A simple positive semi-definite, heteroskedasticity consistent covariance matrix, *Econometrica* **55**, 703–708.

Pastorello, Sergio, Valentin Patilea, and Eric Renault, 2003, Iterative and recursive estimation in structural nonadaptive models, *Journal of Business and Economic Statistics* **21**, 449–482.

Pearson, Neil and T.S. Sun, 1994, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance* **49**, 1279–1303.

Piazzesi, Monika, 2005, Bond yields and the Federal Reserve, *Journal of Political Economy* **113**, 311–344.

Piazzesi, Monika, 2010, Affine term structure models, *Handbook of Financial Econometrics*, Vol. 1, Yacine Aït-Sahalia and Lars P. Hansen, eds., Amsterdam: Elsevier.

Poole, William and Robert Rasche, 2000, Perfecting the market's knowledge of monetary policy, *Journal of Financial Services Research* **18**, 255–298.

Poole, William and Robert Rasche, 2003, The impact of changes in FOMC disclosure practices on the transparency of monetary policy: Are markets and the FOMC better "synched?" *Review of the Federal Reserve Bank of St. Louis* **85**, 1–10.

Renault, Eric, 1997, Econometric models of option pricing errors, *Advances in Economics and Econometrics*, Seventh World Congress, David Kreps and Ken Wallis, eds., New York: Cambridge University Press.

Rudebusch, Glenn D. and Tao Wu, 2008, A macro-finance model of the term structure, monetary policy and the economy, *Economic Journal* **118**, 906–926.

Sims, Christopher A., 2003, Comment (On Iterative and recursive estimation in structural non-adaptive models), *Journal of Business and Economic Statistics* **21**, 500-503.

Stanton, Richard, 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* **52**, 1973–2002.

Swanson, Eric T. and John C. Williams, 2014, Measuring the effect of the zero lower bound on medium- and longer-term interest rates, *American Economic Review* **104**, 3154–3185.

Tang, Huarong and Yihong Xia, 2007, An international examination of affine term structure models and the expectations hypothesis, *Journal of Financial and Quantitative Analysis* **42**, 41-80.

West, Kenneth D., 1996, Asymptotic inference about predictive ability, *Econometrica* **64**, 1067–1084.

Table I

Papers That Estimate and Test Affine Term Structure Models

	No Auxiliary Error Model		Auxiliary Error Model Included			
	Likelihood Based	Quasi-Maximum Likelihood	Likelihood Based	Quasi Maximum Likelihood	Method of Moments	Simulated Maximum Likelihood
OUT-OF-SAMPLE TESTS						
Encompassing Regression	**CP**					
Pricing Errors	**CP**			Duffee (2002) Tang & Xia(2007)		
Density Forecast				Egorov, Hong, & Li (2006)		
Orthogonality Conditions	**CP**			Duffee (2002)		
IN-SAMPLE TESTS						
Fama-Gibbons Regression	PS		Duffie & Singleton (1997)			
Pricing Errors	PS		Bams & Schotman (2003) Chen & Scott (1993) Duffie & Singleton (1997) Lamoureux & Witte (2002) CFK	DeJong (2000) Tang & Xia (2007)	Brown & Dybvig (1986)	Dai & Singleton (2000)
Change in Rates			CFK			
Density Test		Hong & Li (2005)		Egorov, Hong, & Li (2006)		
Nested Likelihood	PS			CFK	DeJong (2000)	
Non-Nested Likelihood			Chen & Scott (1993)			
Conditional Moments						Dai & Singleton (2000)
Unconditional Moments					Gibbons & Ramaswamy(1993)	
Structural Stability	PS				Gibbons & Ramaswamy(1993)	
Eye-Balling Goodness-of-Fit	PS			Egorov, Hong, & Li (2006) DeJong (2000)		Dai & Singleton (2000)

CFK: Cheridito, Filipović, & Kimmel (2007)

PS: Pearson and Sun (1994).

CP: Current Paper.

Table II

Yield Change Encompassing Regressions

Panel A: 3-month yield

Horizon (weeks)	Treasuries Used in Estimating CIR	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate Minus Spot Rate	Aït-Sahalia	Stanton	Bandwidth
1	3-month	1.14	0.263***	-0.009***	-0.005	0.008	0.118***	-0.098***	-0.075	1
		(0.43)	(2.02)	(-2.39)	(-0.96)	(0.86)	(5.00)	(-2.59)	(-1.14)	
1	3-month & 25-year	0.70	0.336	-0.008***	-0.005	0.009	0.122***	-0.104***	-0.087	1
		(0.25)	(1.19)	(-2.07)	(-1.04)	(0.94)	(4.69)	(-2.63)	(-1.22)	
4	3-month	9.62	0.190	-0.038***	-0.026*	0.000	0.370***			22
		(1.22)	(0.95)	(-2.98)	(-1.91)	(0.01)	(5.08)			
4	3-month & 25-year	9.52	-0.144	-0.038***	-0.025*	0.001	0.348***			22
		(1.30)	(-0.73)	(-3.33)	(-1.85)	(0.03)	(4.88)			
13	3-month	39.81	0.155	-0.131***	-0.087*	-0.016	0.982***			40
		(1.51)	(0.63)	(-2.94)	(-1.88)	(-0.19)	(3.89)			
13	3-month & 25-year	40.43*	-0.264	-0.133***	-0.078*	-0.019	0.889***			44
		(1.83)	(-1.19)	(-3.71)	(-1.90)	(-0.23)	(3.78)			

Panel B: 6-month yield

								Bandwidth
1	6-month	5.39*	-0.340	-0.009**	-0.009*	-0.033**	-0.019	2
		(1.94)	(-1.11)	(-2.18)	(-1.80)	(-2.46)	(-0.51)	
1	6-month & 5-year	6.13**	-0.187	-0.010**	-0.009*	-0.035***	-0.020	2
		(2.19)	(-0.80)	(-2.45)	(-1.91)	(-2.60)	(-0.56)	
4	6-month	15.68*	-0.050	-0.038**	-0.028	-0.056	0.139	22
		(1.32)	(-0.21)	(-1.99)	(-1.41)	(-1.47)	(-1.37)	
4	6-month & 5-year	17.19	-0.183	-0.041**	-0.028	-0.061*	0.120	22
		(1.52)	(-0.77)	(-2.27)	(-1.39)	(-1.70)	(1.33)	
13	6-month	47.68	0.150	-0.133***	-0.082	-0.055	0.700**	38
		(1.38)	(0.60)	(-2.35)	(-1.35)	(-0.52)	(2.42)	
13	6-month & 5-year	48.11	-0.113	-0.135***	-0.076	-0.060	0.634**	39
		(1.52)	(-0.34)	(-2.66)	(-1.33)	(-0.69)	(2.44)	

Table II (Continued)

Yield Change Encompassing Regressions

Panel C: 5-year yield

Horizon (weeks)	Treasuries Used in Estimating CIR	Intercept	CIR	Level	Slope	Curvature	Forward Rate Minus Spot Rate	Bandwidth
1	5-year	-1.90	0.401	0.000	0.006	0.024*	0.037	2
		(-0.48)	(1.39)	(-0.04)	(0.84)	(1.93)	(1.58)	
1	6-month & 5-year	-2.38	0.278***	0.001	0.006	0.025**	0.040*	2
		(-0.61)	(3.68)	(0.09)	(0.79)	(2.02)	(1.70)	
4	5-year	-0.22	0.012	-0.012	0.007	0.086**	0.138	25
		(-0.01)	(0.02)	(-0.42)	(0.24)	(2.02)	(1.51)	
4	6-month & 5-year	-1.09	0.222	-0.010	0.007	0.089**	0.146	25
		(-0.06)	(1.28)	(-0.36)	(0.22)	(1.46)	(-0.91)	
13	5-year	1.54	-1.701	-0.044	-0.015	0.249**	0.339*	60
		(0.03)	(-1.60)	(-0.60)	(-0.20)	(2.58)	(1.70)	
13	6-month & 5-year	11.29	-0.073	-0.053	0.002	0.255***	1.66*	62
		(0.29)	(-0.21)	(-0.87)	(0.02)	(2.70)	(1.66)	
Panel D: 15-year yield								
1	15-year	2.10	-1.673	-0.005	-0.004	0.894	0.134	3
		(0.59)	(-2.68)	(-0.92)	(-0.55)	(0.89)	(0.13)	
1	5-year & 15-year	1.96	0.226	-0.004	-0.002	0.010	0.001	3
		(0.55)	(1.03)	(-0.67)	(-0.33)	(0.90)	(0.03)	
4	15-year	11.96	-2.81	-0.031	-0.028	0.035	0.027	20
		(0.80)	(-2.38)	(-1.23)	(-0.95)	(0.89)	(0.36)	
4	5-year & 15 year	13.09	-0.200	-0.025	-0.022	0.034	0.101	20
		(0.90)	(-0.66)	(-1.09)	(-0.79)	(0.88)	(0.14)	
13	15 year	39.04	-3.640	-0.104	-0.101	0.074	-0.003	41
		(0.97)	(-3.09)	(-1.48)	(-1.27)	(1.00)	(-0.02)	
13	5-year & 15 year	46.90	-0.590	-0.087	-0.082	0.084	-0.052	42
		(1.24)	(-0.98)	(-1.43)	(-1.17)	(1.13)	(-0.36)	
Panel E: 25-year yield								
1	25-year	2.65	-0.726	-0.005	-0.004	0.003	-0.011	1
		(0.89)	(-1.23)	(-0.97)	(-0.77)	(0.28)	(-0.60)	
1	5-year & 25-year	2.87	-0.652	-0.005	-0.005	0.002	-0.013	1
		(0.96)	(-1.00)	(-0.98)	(-0.80)	(0.23)	(-0.73)	
4	25-year	6.57	-1.691	-0.018	-0.014	0.036	0.012	22
		(0.51)	(-1.35)	(-0.85)	(-0.59)	(1.08)	(0.18)	
4	5-year & 25-year	8.24	-1.548	-0.018	-0.015	0.034	-0.007	21
		(0.63)	(-1.77)	(-0.81)	(-0.61)	(0.99)	(-0.11)	
13	25-year	16.27	-2.378	-0.054	-0.045	0.082	-0.021	48
		(0.50)	(-1.67)	(-0.96)	(-0.73)	(1.36)	(-0.15)	
13	5-year & 25-year	26.39	-1.164	-0.051	-0.049	0.070	-0.107	48
		(0.78)	(-1.16)	(-0.91)	(-0.79)	(1.15)	(-0.82)	

Table II (Continued)

Yield Change Encompassing Regressions

This table contains encompassing regressions projecting changes in market yields on 3-month, 6-month Treasury bills and approximate 5-year, 15-year and 25-year STRIPS, at the 1-, 4-, and 13-week horizons onto exogenous forecasts.

$$
\begin{aligned}
R^i_{t+\tau} - R^i_t =& \alpha + \beta_{CIR}(\hat{R}^{i,CIR}_{t+\tau} - R^i_t) + \beta_m(\hat{R}^{90,m}_{t+\tau} - R^i_t) \\
&+ \beta_l l_t + \beta_s s_t + \beta_c c_t + \beta_f \hat{f}_t + \epsilon_t
\end{aligned}
$$

Where: R^i_t is the yield on an i-period 0-coupon Treasury security at time t; $\hat{R}^{i,CIR}_{t+\tau}$ is the forecast of the i-period 0-coupon Treasury security at time t+τ, from the CIR model, conditional on information available at time t; l_t is the level of the yield curve at time t, (i.e., the 3-month yield); s_t is the slope of the yield curve at time t, (i.e., the spread between the 25-year and 3-month yields); c_t is the curvature of the yield curve at time t, (i.e., the difference between the spread between the 25-year and 3-month yields and the squared 5-year yield); $\hat{f}_t$ is the difference between the 3-month forward rate, three months hence on date t and the 3-month spot rate on date t; and $\hat{R}^{90,m}_{t+\tau}$ is the forecast of the 3-month yield at time t+τ, from model $m-$ the nonparametric models of Aït-Sahalia and Stanton for the 1-week ahead forecast.

At each forecast horizon, τ, two separate encompassing regressions are reported. Both regressions include the level, slope, and curvature of the yield curve and the difference between the 3-month forward rate, three months hence and the 3-month spot rate, as independent variables. The first regression adds the forecast yield change from a one-factor CIR model, while the second regression includes the forecast change from a two-factor CIR model. For the 3-month yield at the one week horizon both regressions also include the forecast change predicted by Aït Sahalia's (1996) model and Stanton's (1997) model. When estimating the CIR model, we always include the yield being forecast. For the two-factor CIR model, we report only regression results for the combination of yields that results in the highest t-statistic on the two-factor CIR model, at the 1-week horizon. Thus, the t-statistic on the coefficient on the two-factor CIR forecast should be thought of as the maximal t-statistic among the 4 combinations that could be used in estimation, meaning that standard tables should not be used for statistical inference. Statistical significance tests are one-tail (i.e., > 0) for CIR forecasts, and two-tail (i.e., $\neq 0$) for all other exogenous variables. We denote statistical significance at the 10%, 5%, and 1% levels as *, **, and ***, respectively. The bandwidth is the lag length used in calculating the Newey-West (1987) autocovariance matrix. The yields are forecast beginning on 24 March 1994 and ending on 28 June 2007 (693 weeks). We use a rolling sample with 250 weekly observations to estimate the model parameters.

Table III

Yield Spread Change Encompassing Regressions

Panel A: Forecast of 6-month / 3-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	4.04	0.052	0.001	-0.003	-0.041***	-0.137***	5
	(1.42)	(0.10)	(0.15)	(-0.67)	(-3.19)	(-4.09)	
4	6.19	0.108	0.000	-0.002	-0.059***	-0.217***	14
	(0.91)	(0.40)	(0.02)	(-0.22)	(-2.89)	(-4.91)	
13	5.88	0.041	0.000	0.009	-0.037	-0.282***	21
	(0.44)	(0.19)	(0.01)	(0.38)	(-1.12)	(-4.20)	

Panel B: Forecast of 5-year / 3-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	-3.50	0.106	0.010	0.012	0.016	-0.087***	1
	(-0.95)	(1.34)	(1.59)	(1.78)	(1.28)	(-3.23)	
4	-9.47	0.000	0.025	0.033	0.085**	-0.223**	24
	(-0.62)	(0.00)	(1.04)	(1.26)	(1.99)	(-2.36)	
13	-44.21	-0.503	0.106**	0.099*	0.323***	-0.491***	53
	(-1.51)	(-2.98)	(2.21)	(1.74)	(4.76)	(-2.62)	

Panel C: Forecast of 15-year / 3-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	-0.16	0.235	0.007	0.006	0.003	-0.127***	3
	(-0.04)	(1.59)	(1.13)	(0.79)	(0.20)	(-3.63)	
4	2.61	-0.148	0.015	0.005	0.038	-0.337***	21
	(0.18)	(-0.85)	(0.68)	(0.20)	(0.85)	(-3.37)	
13	4.93	-0.424	0.054	-0.004	0.125	-0.899***	37
	(0.13)	(-1.54)	(0.85)	(-0.06)	(1.43)	(-3.72)	

Table III (Continued)

Yield Spread Change Encompassing Regressions

Panel D: Forecast of 25-year / 3-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	1.39	-0.125	0.005**	0.001	-0.006	-0.135***	1
	(0.38)	(-0.53)	(0.96)	(0.19)	(-0.48)	(-4.00)	
4	0.32	-0.322	0.021	0.006	0.026	-0.349***	22
	(0.03)	(-1.27)	(1.12)	(0.27)	(0.63)	(-3.23)	
13	-8.52	-0.284	0.081	0.021	0.071	-1.015***	41
	(-0.30)	(-1.16)	(1.64)	(0.44)	(0.82)	(-3.49)	

Panel E: Forecast of 5-year / 6-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	-7.75**	0.193***	0.009*	0.016***	0.058***	0.049	2
	(-2.45)	(2.44)	(1.84)	(2.76)	(4.04)	(1.46)	
4	16.16	0.001	0.026	0.036*	0.143***	-0.004	23
	(-1.25)	(0.01)	(1.22)	(1.74)	(3.16)	(-0.04)	
13	-39.86	-0.322	0.088*	0.078	0.328***	-0.257	42
	(-1.48)	(-1.95)	(1.95)	(1.59)	(4.10)	(-1.27)	

Panel F: Forecast of 15-year / 6-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	-4.07	0.142	0.006	0.008	0.044***	0.012	4
	(-1.13)	(0.71)	(1.08)	(1.29)	(2.77)	(0.32)	
4	-3.32	-0.026	0.015	0.005	0.096*	-0.114	20
	(0.18)	(-0.99)	(1.86)	(-2.51)	(3.65)	(4.79)	
13	-3.21	-0.365	0.057	-0.007	0.158	-0.642***	32
	(-0.09)	(-1.03)	(0.94)	(-0.11)	(1.63)	(-2.68)	

Panel G: Forecast of 25-year / 6-month Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	-2.45	-0.211	0.004	0.034	0.034**	0.005	2
	(-0.74)	(-0.75)	(0.87)	(0.59)	(2.35)	(0.12)	
4	-3.47	-0.453	0.018	0.002	0.078	-0.124	22
	(-0.29)	(-1.52)	(0.93)	(0.09)	(1.52)	(-1.08)	
13	-10.58	-0.210	0.075	0.008	0.096	-0.758**	36
	(-0.36)	(-0.76)	(1.44)	(0.17)	(0.95)	(-2.58)	

Table III (Continued)

Yield Spread Change Encompassing Regressions

Panel H: Forecast of 15-year / 5-year Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	3.70*	0.136	-0.003	-0.008**	-0.013	-0.038**	6
	(1.90)	(0.98)	(-1.17)	(-2.15)	(-1.54)	(-1.99)	
4	8.82	0.060	-0.004	-0.020*	-0.040	-0.126**	15
	(1.43)	(0.39)	(-0.48)	(-1.93)	(-1.60)	(-2.43)	
13	18.49	0.087	0.001	-0.047*	-0.131**	-0.386***	29
	(1.19)	(0.37)	(0.02)	(-1.93)	(-2.20)	(-3.01)	

Panel I: Forecast of 25-year / 5-year Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	5.16**	-0.361	-0.005	-0.011**	-0.023**	-0.052**	3
	(2.03)	(-1.31)	(-1.22)	(-2.28)	(-2.37)	(-2.51)	
4	10.96	-0.252	-0.005	-0.025*	-0.064**	-0.163***	20
	(1.24)	(-0.54)	(0.37)	(-1.74)	(-2.32)	(-2.73)	
13	21.43	-0.890	0.000	-0.053	-0.203***	-0.455***	42
	(1.08)	(-1.54)	(0.01)	(-1.62)	(-3.13)	(-3.00)	

Panel J: Forecast of 25-year / 15-year Spread

Horizon (Weeks)	Intercept (Basis Points)	CIR	Level	Slope	Curvature	Forward Rate minus Spot Rate	Bandwidth
1	1.84	1.174	-0.002	-0.004*	-0.012***	-0.016**	9
	(1.44)	(1.33)	(-0.94)	(-1.70)	(-3.08)	(-2.46)	
4	3.21	1.634**	-0.002	-0.010	-0.030**	-0.037*	11
	(0.76)	(2.22)	(-0.36)	(-1.22)	(-2.49)	(-1.73)	
13	6.39	0.888	-0.005	-0.020	-0.095***	-0.091	19
	(0.55)	(0.86)	(-0.26)	(-0.90)	(-3.01)	(-1.52)	

This table contains the encompassing regressions for changes in the ten pair-wise term spreads amongst the five yields in our sample at the 1-, 4-, and 13-week horizons.

$$\begin{aligned}(R^i_{t+\tau} - R^j_{t+\tau}) - (R^i_t - R^j_t) =& \alpha + \beta_{CIR}\left[(\hat{R}^{i,CIR}_{t+\tau} - \hat{R}^{j,CIR}_{t+\tau}) - (R^i_t - R^j_t)\right] \\ &+ \beta_l l_t + \beta_s s_t + \beta_c c_t + \beta_f \hat{f}_t + \epsilon_t\end{aligned}$$

Where: R^i_t is the yield on an i-period 0-coupon Treasury security at time t; $\hat{R}^{i,CIR}_{t+\tau}$ is the forecast of the i-period 0-coupon Treasury security at time t+τ, from the CIR model, conditional on information available at time t; l_t is the level of the yield curve at time t, (i.e., the 3-month yield); s_t is the slope of the yield curve at time t, (i.e., the spread between the 25-year and 3-month yields); c_t is the curvature of the yield curve at time t, (i.e., the difference between the spread between the 25-year and 3-month yields and the squared 5-year yield); and $\hat{f}_t$ is the difference between the 3-month forward rate, three months hence and the 3-month spot rate on date t. The two yields that define the term spread being forecast are used to estimate the CIR model. The yields are forecast beginning on 24 March 1994 and ending on 28 June 2007 (693 weeks). We use a rolling sample with 250 weekly observations to estimate the model parameters. Statistical significance tests are one-tail (i.e., > 0) for CIR forecasts, and two-tail (i.e., $\neq 0$) for all other exogenous variables. The bandwidth is the lag length used in calculating the Newey-West (1987) autocovariance matrix, obtained using Andrews' (1991) procedure.

Table IV

Testing Unconditional and Conditional Predictive Ability
for Changes in U.S. Treasury Yields
in Basis Points

Panel A: Forecast of 3-month yield

	CIR Model estimated using:								
Horizon (weeks)	3-month	3-month & 6-month	3-month & 5-year	3-month & 15-year	3-month & 25-year	Martingale	3-month Forward	Aït Sahalia	Stanton
1	10.087	10.120	10.153	10.143	10.090	**9.799**		13.623	13.245
	(39.1***)	(116.3***)	(60.3***)	(93.6***)	(69.6***)			(760.7***)	(280.6***)
	[37.4***]	[44.4***]	[55.9***]	[79.5***]	[149.8***]			[3.9]	[2002.9***]
4	21.944	22.740	23.055	22.924	22.663	**20.567**			
	(75.4***)	(88.9***)	(78.9***)	(84.1***)	(78.6***)				
	[1037.3***]	[2758.1***]	[2088.3***]	[2357.1***]	[3674.2***]				
13	51.028	53.747	54.391	54.480	53.832	**45.953**	48.086		
	(28.3***)	(52.4***)	(45.6***)	(48.4***)	(39.2***)		(3.3*)		
	[15728.4***]	[16435.9***]	[8702.1***]	[13845.7***]	[18086.3***]		[227.6***]		

Panel B: Forecast of 6-month yield

Horizon (weeks)	6-month	6-month & 3-month	6-month & 5-year	6-month & 15-year	6-month & 25-year	Martingale
1	9.997	10.135	10.056	10.049	10.043	**9.764**
	(106.8***)	(136.7***)	(75.4***)	(109.0***)	(125.5***)	
	[171.4***]	[180.5***]	[57.2***]	[90.8***]	(268.8***)	
4	22.452	23.741	23.266	23.199	23.086	**21.096**
	(48.0***)	(93.1***)	(67.8***)	(71.3***)	(73.3***)	
	[1662.0***]	[1577.2***]	[1728.7***]	[2225.2***]	[3252.2***]	
13	52.048	56.176	54.306	54.806	54.442	**47.811**
	(27.1***)	(59.4***)	(42.9***)	(50.0***)	(43.9***)	
	[26379.8***]	[8268.2***]	[6314.1***]	[15279.9***]	[20123.7***]	

Panel C: Forecast of 5-year yield

Horizon (weeks)	5-year	5-year 3-month	5-year 6-month	5-year 15-year	5-year 25-year	Martingale
1	13.486	13.973	13.847	13.470	13.483	**13.465**
	(0.9)	(83.0***)	(109.4***)	(0.2)	(1.4)	
	[1.9]	[26.1***]	[7.5]	[0.1]	[3.2]	
4	27.989	28.663	28.558	27.960	27.998	**27.819**
	(11.4***)	(13.8***)	(11.1***)	(11.6***)	(16.8***)	
	[24.9***]	[40.2***]	[29.6***]	[49.0***]	[89.5***]	
13	52.804	52.982	52.787	52.702	52.543	**51.703**
	(21.1***)	(12.2***)	(8.8***)	(29.0***)	(18.3***)	
	[669.6***]	[396.2***]	[292.6***]	[585.3***]	[588.9***]	

Table IV (Continued)

Testing Unconditional and Conditional Predictive Ability for Changes in U.S. Treasury Yields

in Basis Points

Panel D: Forecast of 15-year yield

CIR Model estimated using:

Horizon (weeks)	15-year	15-year & 3-month	15-year & 6-month	15-year & 5-year	15-year & 25-year	Martingale
1	12.959	13.084	13.000	13.002	12.930	**12.910**
	(24.9***)	(12.8***)	(4.7**)	(2.1)	(5.6**)	
	[24.6***]	[12.5**]	[4.5]	[2.2]	[6.0]	
4	23.783	24.165	23.989	24.042	23.688	**23.601**
	(24.1***)	(25.4***)	(16.7***)	(19.6***)	(7.9***)	
	[71.4***]	[45.4***]	[28.9***]	[42.7***]	[23.7***]	
13	41.157	40.802	40.899	40.743	40.638	**40.571**
	(10.5***)	(0.6)	(1.6)	(0.6)	(0.2)	
	[268.0***]	[9.5*]	[20.3***]	[8.4]	[3.0]	

Panel E: Forecast of 25-year yield

Horizon (weeks)	25-year	25-year & 3-month	25-year & 6-month	25-year & 5-year	25-year & 15-year	Martingale
1	10.581	10.731	10.673	10.626	10.636	**10.542**
	(11.3***)	(16.6***)	(14.4***)	(33.4***)	(80.5***)	
	[8.5]	[17.9***]	[7.9]	[2.2]	[2.4]	
4	20.724	21.081	20.972	20.889	20.913	**20.593**
	(16.0***)	(64.8***)	(53.8***)	(46.0***)	(58.5***)	
	[33.1***]	[194.0***]	[131.2***]	[140.3***]	[200.3***]	
13	35.468	35.401	**35.099**	35.166	35.300	35.203
	(3.9**)	(4.0**)	(0.0)	(0.0)	(0.7)	
	[51.3***]	[82.5***]	[0.1]	[0.2]	[7.8]	

Table IV (Continued)

Testing Unconditional and Conditional Predictive Ability for Changes in U.S. Treasury Yields

in Basis Points

This table presents the Root Mean Square Error (RMSE) of forecasting zero-coupon yields at the 1-, 4-, and 13-week horizons using the forward rate, martingale model, Aït Sahalia's (1996) model, Stanton's (1997) model, and both one and two-factor CIR models. The one-factor CIR model is always estimated using the same yield being forecast. The two-factor CIR model is estimated in turn by combining the yield being forecast with each of the other four yields in the sample. The out-of-sample period comprises 693 weeks, beginning 24 March 1994, and ending on 28 June 2007. We use a rolling estimation design with (constant) sample size of 250 weekly observations to estimate the model parameters. Inference is conducted using the test of Giacomini and White (2006):

$$T^h_{m,n} = n\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right)' \widehat{\Omega}^{-1}\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right)$$

Where: $\Delta L_{t+\tau}$ is the difference in the squared error of the two models' forecasts of the change in the yield from time t to time t+τ, conditional on the information set at time t. We report this statistic for two different values of h_t: the unconditional test statistic with $h^{'}_t = \{1\}$ in parentheses, and the conditional test statistic with $h^{'}_t = \left\{1, l_t, s_t, c_t, \hat{f}_t\right\}$ in square brackets below the corresponding RMSE. Where: l_t is the level of the yield curve at time t, (i.e., the 3-month yield); s_t is the slope of the yield curve at time t, (i.e., the spread between the 25-year and 3-month yields); c_t is the curvature of the yield curve at time t, (i.e., the difference between the spread between the 25-year and 3-month yields and the squared 5-year yield); and $\hat{f}_t$ is the difference between the 3-month forward rate, three months hence and the 3-month spot rate on date t. The covariance matrix, $\widehat{\Omega}$ is the Newey-West (1987) estimator, with the lag length (or bandwidth) selected using the procedure of Andrews (1991). The null hypothesis is that the expected difference in the forecast methods from the martingale and the *estimated* model is zero, and orthogonal to the state variables in $h^{'}_t$.

We indicate statistical significance of the conditional test of this null hypothesis with *, **, and ***, for significance at the 10%, 5%; and 1% levels, respectively. The lowest RMSE forecast for each yield is in **boldface.**

Table V

Testing Unconditional and Conditional Predictive Ability for Changes in U.S. Treasury Yield Spreads

March 1994 through June 2007

Horizon	**Forecast Change in:** 6-month Minus 3-month		**Forecast Change in:** 5-year Minus 3-month		**Forecast Change in:** 15-year Minus 3-month	
(Weeks)	CIR	Martingale	CIR	Martingale	CIR	Martingale
1	**8.409**	8.414	14.711	**13.731**	15.429	**15.106**
	(0.1)		(118.6***)		(29.9***)	
	[0.1]		[78.9***]		[46.5***]	
4	11.222	**11.201**	29.932	**27.669**	31.036***	**29.058**
	(0.2)		(72.8***)		(130.0***)	
	[0.2]		[408.5***]		[476.6***]	
13	15.428	**14.782**	60.833***	**53.209**	64.199	**57.399**
	(9.9***)		(57.4***)		(58.3***)	
	[97.8***]		[6971.2***]		[5360.8***]	

Horizon	25-year Minus 3-month		5-year Minus 6-month		15-year Minus 6-month	
(Weeks)	CIR	Martingale	CIR	Martingale	CIR	Martingale
1	14.179	**13.758**	12.735	**12.023**	13.831	**13.594**
	(48.9***)		(146.6***)		(9.3***)	
	[67.0***]		[25.5***]		[19.9***]	
4	29.963	**27.894**	26.206	**23.558**	28.147	**26.394**
	(112.3***)		(59.5***)		(74.8***)	
	[894.6***]		[523.8***]		[474.4***]	
13	63.434	**56.356**	52.921	**45.939**	58.739	**53.079**
	(60.6***)		(52.5***)		(40.5***)	
	[11169.4***]		[8834.7***]		[4495.9***]	

Horizon	25-year Minus 6-month		15-year Minus 5-year		25-year Minus 5-year	
(Weeks)	CIR	Martingale	CIR	Martingale	CIR	Martingale
1	12.642	**12.339**	9.862	**9.672**	9.798	**9.669**
	(48.9***)		(5.3**)		(19.1***)	
	[33.2***]		[4.9]		[16.6***]	
4	27.838	**25.796**	16.317	**15.765**	18.212	**17.841**
	(109.5***)		(12.2***)		(70.6***)	
	[1414.9***]		[19.5***]		[217.2***]	
13	59.793	**53.107**	27.136	**26.222**	33.529	**32.315**
	(56.9***)		(40.3***)		(60.0***)	
	[22822.4***]		[206.5***]		[1618.2***]	

Table V (Continued)

Testing Unconditional and Conditional Predictive Ability for Changes in U.S. Treasury Yield Spreads

in Basis Points

March 1994 through June 2007

	Forecast Change in:	
Horizon	25-year Minus 15-year	
(Weeks)	CIR	Martingale
1	**6.992**	7.063
	(0.4)	
	[0.6]	
4	**8.924**	9.279
	(12.0***)	
	[20.0***]	
13	**14.038**	14.160
	(0.2)	
	[7.8]	

This table presents the Root Mean Square Error (RMSE) of forecasting the spreads between yield pairs at 1-, 4-, and 13-week horizons using the independent-bivariate martingale model and the two-factor CIR model. The out-of-sample period comprises 693 weeks, beginning 24 March 1994, and ending on 28 June 2007. We use a rolling design with 250 weekly observations in estimating the model parameters. Inference is conducted using the test of Giacomini and White (2006):

$$T_{m,n}^{h} = n\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right)' \widehat{\Omega}^{-1}\left(n^{-1}\sum_{t=m}^{T-1} h_t \Delta L_{t+\tau}\right)$$

Where: $\Delta L_{t+\tau}$ is the difference in the squared error of the two models' forecasts of the change in the yield spread from time t to time t+τ, conditional on the information set at time t. We report this statistic for two different values of h_t: the unconditional test statistic with $h_t' = \{1\}$ in parentheses, and the conditional test statistic with $h_t' = \left\{1, l_t, s_t, c_t, \hat{f}_t\right\}$ in square brackets below the corresponding RMSE. Where: l_t is the level of the yield curve at time t, (i.e., the 3-month yield); s_t is the slope of the yield curve at time t, (i.e., the spread between the 25-year and 3-month yields); c_t is the curvature of the yield curve at time t, (i.e., the difference between the spread between the 25-year and 3-month yields and the squared 5-year yield); and $\hat{f}_t$ is the difference between the 3-month forward rate, three months hence on date t and the 3-month spot rate on date t. The covariance matrix, $\widehat{\Omega}$ is the Newey-West (1987) estimator, with the lag length (or bandwidth) selected using the procedure of Andrews (1991). The null hypothesis is that the expected difference in the forecast methods from the martingale and the *estimated* model is zero, and orthogonal to the state variables in h_t'.

We indicate statistical significance of the conditional test of this null hypothesis with *, **, and ***, for significance at the 10%, 5%; and 1% levels, respectively. The lowest RMSE forecast for each yield is in **boldface.**

Table VI

Decision Rule Assessment
Significant Effects of State Variables on the CIR Forecast
Minus Martingale Squared Errors

Forecasted Variable	Horizon (Weeks)	Model	Effect of 1 Standard Deviation Increase in: Level	Slope	Curvature	Forward Rate minus Spot
3-mo yield	1	1F CIR	-0.136	-0.138	–	0.074
3-mo yield	4	1F CIR	-1.084	-1.122	–	0.925
3-mo yield	13	1F CIR	-10.376	–	–	7.863
3-mo yield	13	Forward Rate	–	–	–	–
6-mo yield	1	1F CIR	-0.090	-0.104	–	–
6-mo yield	4	1F CIR	–	–	–	–
6-mo yield	13	1F CIR	–	–	–	6.313
5-year yield	1	CIR (5-yr, 15-yr)	–	–	–	–
5-year yield	4	CIR (5-yr, 15-yr)	–	–	–	–
5-year yield	13	CIR (5-yr, 15-yr)	–	–	–	–
15-year yield	1	CIR (15-yr, 25-yr)	–	–	–	–
15-year yield	4	CIR (15-yr, 25-yr)	–	–	–	–
15-year yield	13	CIR (15-yr, 25-yr)	–	–	–	–
25-year yield	1	CIR (25-yr, 6-mo)	–	–	–	–
25-year yield	4	CIR (25-yr, 6-mo)	–	–	–	–
25-year yield	13	CIR (25-yr, 6-mo)	–	–	–	–
6-mo - 3-mo spread	1	2F CIR	–	–	–	–
6-mo - 3-mo spread	4	2F CIR	–	–	–	–
6-mo - 3-mo spread	13	2F CIR	–	–	–	–
5-yr - 6-mo spread	1	2F CIR	–	–	–	–
5-yr - 6-mo spread	4	2F CIR	–	–	–	–
5-yr - 6-mo spread	13	2F CIR	–	–	–	–
15-yr - 5-yr spread	1	2F CIR	–	–	–	–
15-yr - 5-yr spread	4	2F CIR	–	–	–	–
15-yr - 5-yr spread	13	2F CIR	-0.736	–	–	–
25-yr - 15-yr spread	1	2F CIR	–	–	–	–
25-yr - 15-yr spread	4	2F CIR	–	–	–	–
25-yr - 15-yr spread	13	2F CIR	–	–	–	–

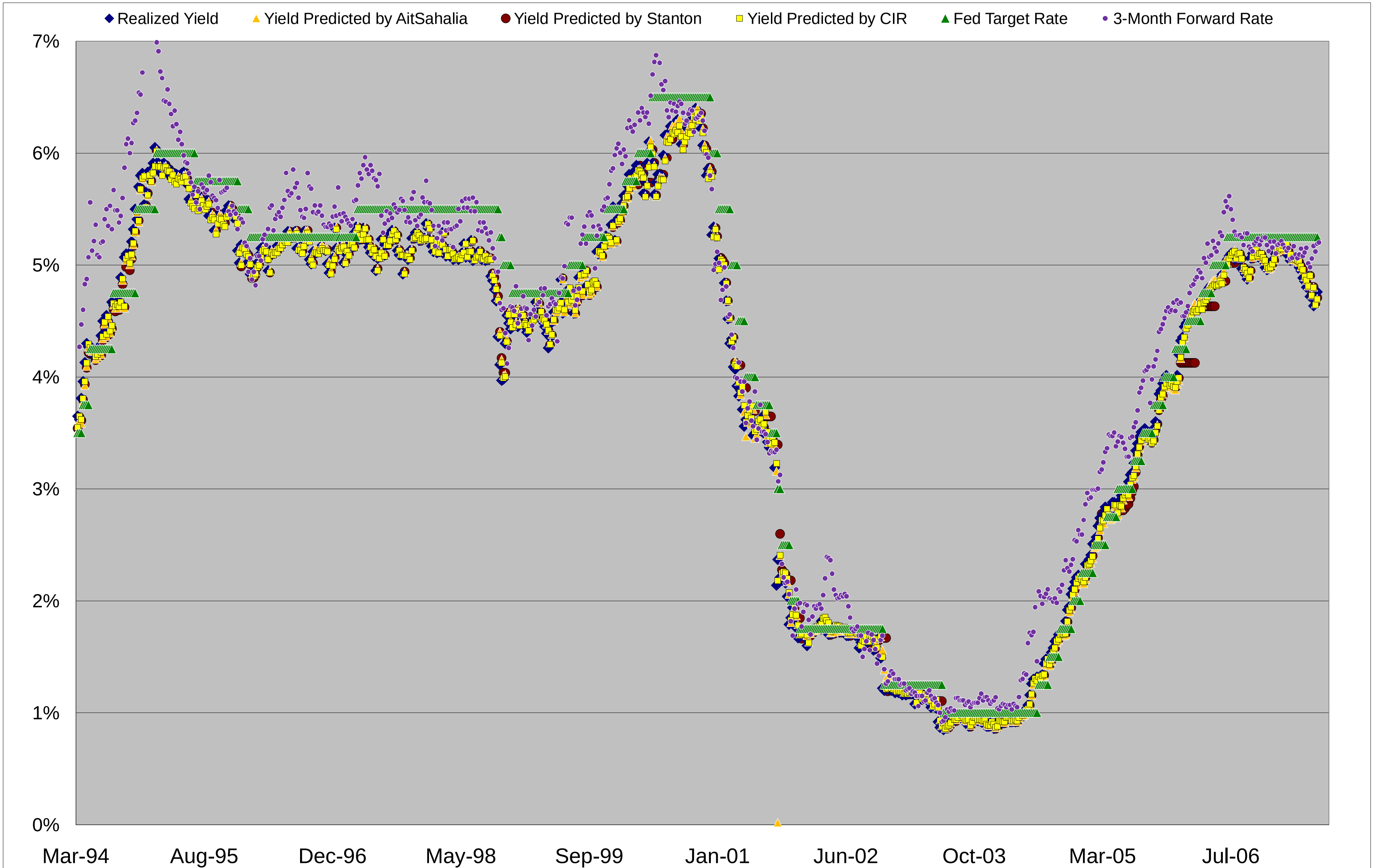
This table reports the effects of a one standard deviation increase in a state variable on the difference in the squared errors between the indicated model and the martingale for statistically significant cases. That is, we estimate the linear regression:

$$\Delta L_{t+\tau} = \gamma_0 + \gamma_1 l_t + \gamma_2 s_t + \gamma_3 c_t + \gamma_4 \hat{f}_t + \epsilon_t \tag{A.5}$$

Where: $\Delta L_{t+\tau}$ is the difference in the squared error of the two models' forecasts of the change in the yield from time t to time t+τ, conditional on the information set at time t; l_t is the level of the yield curve at time t, (i.e., the 3-month yield); s_t is the slope of the yield curve at time t, (i.e., the spread between the 25-year and 3-month yields); c_t is the curvature of the yield curve at time t, (i.e., the difference between the spread between the 25-year and 3-month yields and the squared 5-year yield); and $\hat{f}_t$ is the difference between the 3-month forward rate, three months hence and the 3-month spot rate on date t.

FIGURE 1

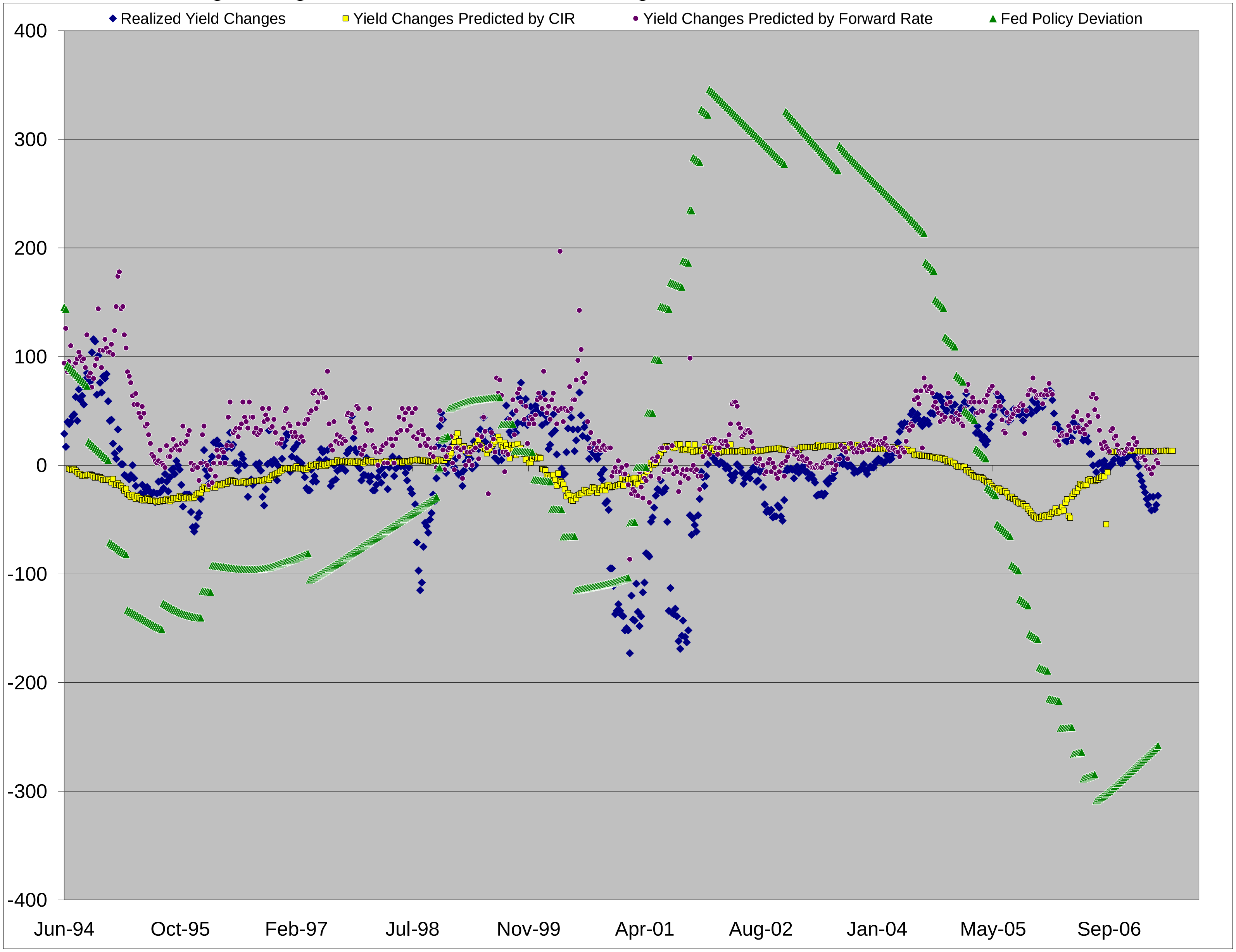
Forecasting the 3-Month Yield Using a 2 Factor CIR Model at a 1 Week Horizon



This figure plots the 1-week ahead 3-month yield against the predicted 1-week ahead 3-month yield from a two factor Cox, Ingersoll, and Ross (1986) term structure model which is estimated using the 3-month and 25-year yields. Also plotted are the predicted yields from the nonparametric models of Ait-Sahalia (1996) and Stanton (1997), as well as the Federal Reserve's target Federal Funds rate at the time the forecast is constructed. The sample period begins on March 31, 1994 and extends until June 28, 2007.

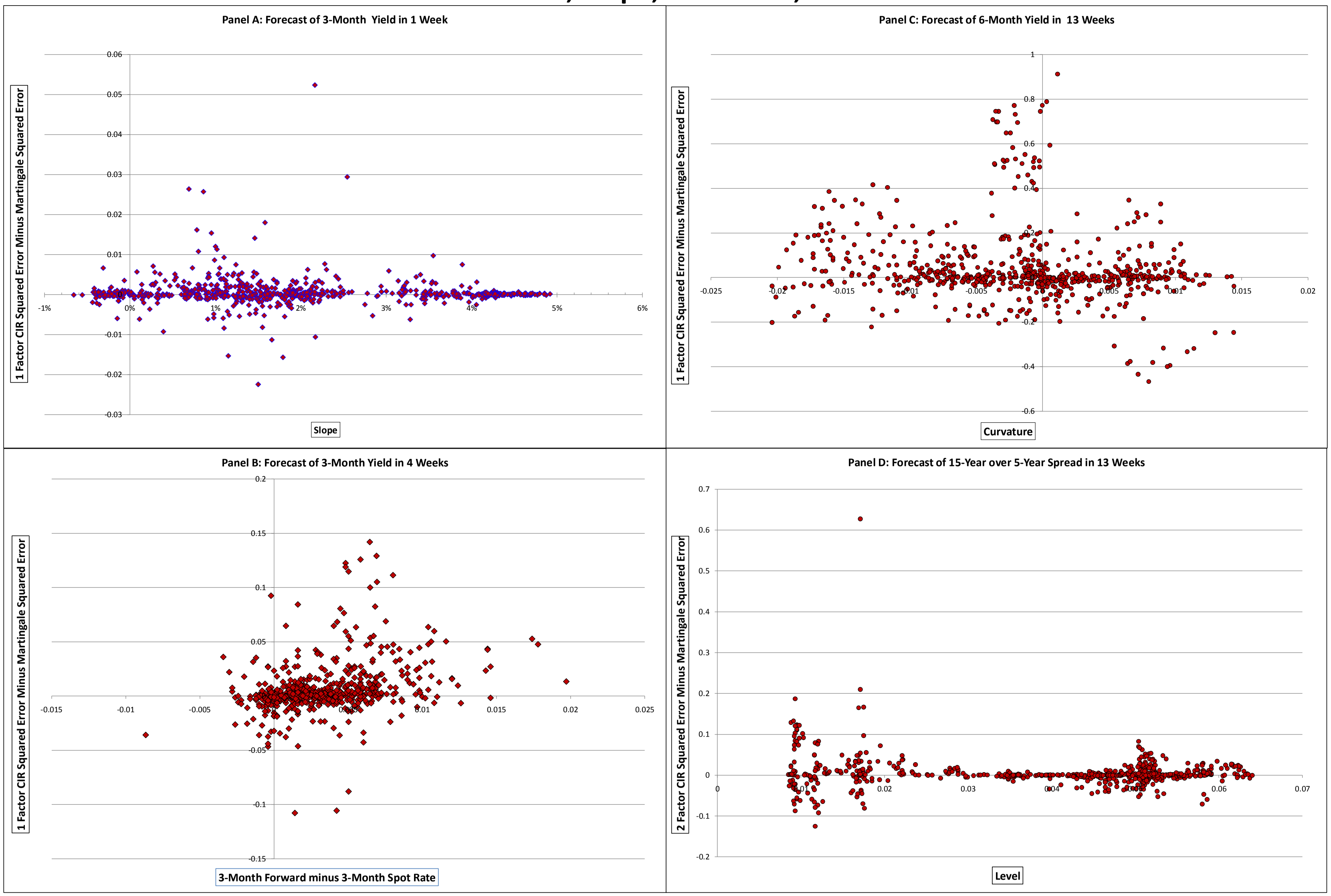
FIGURE 2

Forecasting Changes in the 3-Month Yield Using a 2 Factor CIR Model at a 13 Week Horizon



This figure plots the realized 13-week change in the 3-month yield against the predicted 13-week yield change from a two factor Cox, Ingersoll, and Ross (1986) term structure model estimated using the 3-month and 25-year yields. The change implied by the 13-week forward rate is included, as well as the difference between the average Federal Reserve's target Federal Funds rate over the previous 250 weeks and the target Fed Funds rate at the time of the forecast. The sample period begins on March 31, 1994 and extends until June 28, 2007.

Figure 3
Difference in the Forecast Squared Error of CIR and the Martingale Models as a function of Level, Slope, Curvature, and Forward Rate



These figures plot the difference in the squared forecast error of the CIR and martingale models as a function of the level, slope, and curvature of the yield curve, as well as the forward rate minus the level. Level is defined as the 3-month yield on the day the forecast is constructed, while slope is the difference in the 25-Year yield and the 3-Month yield, and curvature is the 25-Year yield plus the 3-Month yield minus twice the 5-Year yield. The forward rate is the implied 3-month rate in 3 months.