Three Essays on the Term Structure of Interest Rates

Luiz Paulo Fichtner

Nova School of Business and Economics

Universidade Nova de Lisboa

A thesis submitted for the degree of

Doctor of Philosophy

March 2013

Abstract

This dissertation comprises of three essays about the term structure of interest rates. The two first chapters are joined works with my PhD thesis advisor, Pedro Santa-Clara. More than being studies on the same underlying asset (bonds), what binds these essays together is the use of simple ideas to bring light to some problems in the literature. In the first essay I show that models proposed in the literature to explain bond excess returns fail to perform out of sample. Instead of regressing returns on a set of explanatory variables, I forecast bond yields and replace them directly in the bond excess return definition. An investor who used a simple random walk on yields would have predicted bond excess returns with out-of-sample R-squares of up to 15%, while a dynamic Nelson-Siegel approach would have produced out-of-sample R-squares of up to 30%. On the second and third essays I evaluate the performance of a two-factor Cox et al. (1985a,b) model estimation using a state-space framework, while changing the weights in the joint likelihood function. Using EURIBOR zero-coupon yields I show that giving more weight to the likelihood of pricing errors improves the fitting and forecasting of EURIBOR yields, while giving more weight to the likelihood of short rate factor dynamics improves interest rate derivative pricing at the expense of the first. On the third essay I further explore this idea using Monte Carlo simulations. I show that the bias found in Ball and Torous (1996) and Phillips and Yu (2005) can be minimised by giving more weight to the likelihood of pricing errors. As a consequence, fitting the yield curve and bond option pricing performance of the model greatly improves. When noise is added to model-implied bond yields, this still holds true for the one-factor model. However, the bond versus bond option pricing tradeoff is observed for the two-factor model.

Contents

List of Figures

List of Tables

1

Out-of-Sample Predictability of Bond Returns

1.1 Introduction

Significant predictability of bond returns is a much established fact in the literature. Cochrane and Piazzesi (2005) find that a single tent-shaped linear combination of forward rates predicts excess returns on one- to five-year maturity bonds with R-square up to 44%. They strengthen the results by Fama and Bliss (1987) and Campbell and Shiller (1991) against the Expectation Hypothesis of the Term Structure (EH). Based on a modified version of the EH, in which expected bond excess returns are unforecastable, Fama and Bliss test whether forward rates have information about expected returns on bonds with different maturities. They find that the spread between the n-year forward rate and the one-year yield predicts the one-year excess return of the n-year bond, with R-squares up to 18%. Using a comparable approach, Campbell and Shiller (1991) find similar results forecasting yield changes with yield spreads. Following the bond return predictability literature, Ilmanen (1995) forecasts out-of-sample bond excess returns for different countries using the term spread, real yields, the inverse relative wealth, and bond return betas. He finds that these variables predict monthly bond excess returns with out-of-sample R-squares up to 12%.[1] Greenwood and Vayanos (2008) find that, consistent with a preferred habitat model of the term structure, where clienteles with

[1] Ilmanen (1995) also uses international weighted averages versions of these variables. But due to data availability for our samples, and because we found no impact of these variables for the data available, we chose to use only US variables.

strong preferences for specific maturities trade with arbitrageurs, the supply of long- relative to short-term bonds positively predicts long-term excess returns even after controlling for the term spread and Cochrane and Piazzesi's single factor.

In this paper we assess the out-of-sample predictive power of the methods proposed by these authors for one-year holding period excess return regressions. We use Fama-Bliss monthly one- to five-year zero coupon bond prices (available from CRSP), from January 1965 to December 2010. On this dataset, we replicate the regressions of Fama and Bliss (1987), Ilmanen (1995), Cochrane and Piazzesi (2005), and Greenwood and Vayanos (2008), and forecast one-year holding period excess returns using only data available at the time of the forecast.[1] This is the only approach that could have been implemented to time the market in real time. Judging by the R-square values, the models do well in explaining bond excess returns *in sample*. The Fama and Bliss regression produces R-squares of up to 15%, Ilmanen and Greenwood and Vayanos regressions produce R-squares of up to 31%, and Cochrane and Piazzesi's regression reach R-squares of up to 33%.

However, hight in-sample R-squares do not imply good out-of-sample performance. Using the metric proposed by Goyal and Welch (2003), we find that all methods, except for Fama and Bliss, produce negative out-of-sample R-squares. Fama and Bliss' method produces out-of-sample R-squares of up to 8.5%. Ilmanen's regressions produce negative out-of-sample R-squares for the two- and five-year bond returns, while small, but positive out-of-sample R-squares for the three- and four-year bond returns. Cochrane and Piazzesi's single factor regression produces out-of-sample R-squares of -8.5%. Greenwood and Vayanos' regressions have the worst performance, with out-of-sample R-squares of -11%. Since the regressions try to minimize squared errors, they tend to overfit in sample.[2] Sampling noise in the data make regression coefficients not robust out of sample. Furthermore, the data is also prone to pure measurement error. In practice, discount bond prices at exactly the desired maturities are not observed. Instead, they must be

[1] Campbell and Shiller (1991) perform regressions based on yield spreads and these behave similarly to Fama and Bliss' regression. Thus, for simplicity, we only reproduce Fama and Bliss (1987).

[2] Ashley (2006) shows that the unbiased forecast is no longer squared-error optimal in this setting. Instead, the minimum-MSE forecast is shown to be a shrinkage of the unbiased forecast toward zero. However, applying shrinkage to our regressions did not alter significantly the results. We therefore chose not to report shrinked out-of-sample R-squares.

estimated from observed bond prices. If excess return components are also used as right-hand-side variables, as in the case of Fama and Bliss's forward spread and Cochrane and Piazzesi's single factor, measurement error will induce spurious common movement in left- and right-hand-side variables. Although Cochrane and Piazzesi (2005) argue that the results of their single factor approach are not driven by measurement error, our out-of-sample results suggest that this and overfitting make their approach useless to forecast returns in practice.

In this paper we propose a different method for predicting bond excess returns. Instead of regressing returns on a set of explanatory variables, we forecast the relevant return components separately. By definition, the one-year holding period log excess return on a n-year bond at time $t+1$ is the difference between n times the n-maturity bond yield at time t, and the sum of $n-1$ times the $(n-1)$-maturity bond yield at time $t+1$ and the one-year bond yield at time t. It is straightforward to see that, at time t, the $(n-1)$-maturity bond yield at time $t+1$ is the only unknown. We forecast these bond yields to produce bond return forecasts. We use two yield forecasting approaches. First, simply we assume that yields follow a random walk. Substituting the random walk yield forecast in the excess return definition, the excess return forecast at time $t+1$ will be equal to the forward spread at time t. Although very simple, this approach produces out-of-sample R-squares of up to 15%. It also strengthens the idea that overfitting and measurement errors affects bond excess return forecasts based on predictive regressions. Fama and Bliss use the lagged forward spread in their excess return predictions, but assuming that yields follow a random walk is equivalent to taking the forward spread directly as the return forecast. Second, we forecast future yields using a dynamic Nelson Siegel model proposed by Diebold and Li (2006). This approach is highly tractable, and produces reasonable out-of-sample yield forecasts.[1] Out-of-sample R-squares reach up to 24% with the dynamic Nelson Siegel approach using the Fama-Bliss dataset.

To verify the robustness of our findings, we test the out-of-sample predictive power of the Fama and Bliss (1987) and dynamic Nelson Siegel methods on a second dataset, namely Gurkaynak et al. (2006) monthly one- to twenty-year zero coupon bond prices (available from the Federal Reserve Board), from July 1982 to December 2010. This

[1] The out-of-sample R-square measure depends exclusively on the historical sample mean and the forecasting errors. Since our approach is to forecast yields separately and plug them back into the excess return equation, the results are independent of the method used to forecast yields.

dataset enables us to evaluate the models using a more representative cross-section of the term structure of bonds, albeit for a shorter time series. When the Fama and Bliss model is estimated using this smaller sample, regression coefficients are no longer significant and out-of-sample R-squares are negative for all maturities. The three- and five-year bond excess return regressions that share a common sample between both datasets produce out-of-sample R-squares of -33% and -29%, respectively, in comparison to 7% and 8.5% in the Fama-Bliss dataset. On the other hand, the dynamic Nelson Siegel approach benefits from using a larger cross-section of bond yields in its estimation and produces out-of-sample R-squares of up to 30%.

The paper proceeds as follows. Section 3.2 discusses the notation, the methodology of the traditional predictive regressions and the performance measure. Section 1.3 describes the data, the main empirical results and the robustness check. Section 3.4 concludes.

1.2 Forecasting Returns with Predictive Regressions

The central goal of this paper is to assess the performance of existing Bond return predictability models, and compare them with our yield forecasting method. For this purpose, we need to evaluate models both by their in-sample and their out-of-sample performance. We begin this section by describing the notation and predictive regression methodology that we are using throughout the paper.

1.2.1 Notation

We use the same notation as CP and denote log bond prices as:

$$p_t^{(n)} = \text{ log price of } n\text{-year discount bond at time.} \tag{1.1}$$

Parenthesis are used to distinguish maturity from exponentiation in the superscript. The log yield is:

$$y_t^{(n)} \equiv -\frac{1}{n} p_t^{(n)}, \tag{1.2}$$

and the log forward rate at time t for loans between time $t+n-1$ and $t+n$ is:

$$f_t^{(n)} \equiv p_t^{(n-1)} - p_t^{(n)}. \tag{1.3}$$

The log holding period excess return from buying an n-year bond at time t and selling it as an $n-1$ year bond at time $t+1$ is:

$$r_{t+1}^{(n)} \equiv p_{t+1}^{(n-1)} - p_t^{(n)} - y_t^{(1)}. \tag{1.4}$$

Bold letters denote vectors. When used as right-hand-side variables, these vectors include an intercept. The traditional predictive regression methodology regresses excess returns on lagged predictors $\mathbf{X}_{t-1}$:

$$\mathbf{r}_t = \beta^\intercal \mathbf{X}_{t-1} + \varepsilon_t.$$

We generate out-of-sample forecasts of the vector of log bond excess returns using a sequence of expanding windows, starting with an initial sample of five years. We use a subsample $t = 1, ..., s$ of the entire sample of T observations and estimate the model. The forecast of the one-year excess log return at time $s+1$ is the estimated coefficient (denoted with hat) times the contemporaneous value of the predictor:

$$\widehat{\mathbf{r}}_{s+1} = \widehat{\beta}^\intercal \mathbf{X}_s.$$

1.2.2 Yield Forecasting Approach

We propose a new method for predicting log bond excess returns. We rewrite the one-year bond excess return definition in (1.4) in terms of yields:

$$r_{t+1}^{(n)} = (n-1) y_{t+1}^{(n-1)} - n y_t^{(n)} - y_t^{(1)}. \tag{1.5}$$

It is straightforward to see from the equation above that, at time t, $y_{t+1}^{(n-1)}$ is the only unknown. Our approach is to forecast $y_{t+1}^{(n-1)}$ and replace it in the excess return definition to produce our excess return forecast. We use two methods to forecast yields. First, we assume that yields follow a random walk. In this way, we do not need to estimate a yield forecast for every period in our out-of-sample procedure. Assuming a random walk for bond yields, we can write our excess return forecast as

$$\widehat{r}_{t+1}^{(n)} = (n-1) y_t^{(n-1)} - n y_t^{(n)} - y_t^{(1)}.$$

Using the definitions in (1.2) and (1.3), we rewrite the equation above as:

$$\widehat{r}_{t+1}^{(n)} = f_t^{(n)} - y_t^{(1)}. \tag{1.6}$$

That is, the random walk excess return forecast equals the forward spread. In essence, we are using the same component as Fama and Bliss (1987), except that we do not run a regression, but instead use the forward spread directly as our return forecast. Second, we forecast yields using a dynamic Nelson Siegel model proposed by Diebold and Li (2006). Diebold and Li extend the the Siegel and Nelson (1988) model to an out-of-sample forecasting exercise. The Nelson-Siegel approach assumes that forward rates (and thus yields) can be characterised by a continuous function with only four parameters, as a polynomial times an exponential decay term.[1] They fit the yield curve using the three factor model

$$y_t^{(n)} = L_t + S_t \left(\frac{1 - e^{-\lambda_t n}}{\lambda_t n} \right) + C_t \left(\frac{1 - e^{-\lambda_t n}}{\lambda_t n} - e^{-\lambda_t n} \right). \tag{1.7}$$

Where λ_t governs the exponential decay terms. To estimate their model, they fix the parameter that governs the exponential decay rates at 0.0609. This value maximizes the loading on the curvature factor at 30 months. By fixing λ_t, they are able to estimate the remaining parameters by ordinary least squares. Diebold and Li interpret the other three parameters L_t, S_t and C_t as latent dynamic factors, and show that they relate to the level, slope and curvature of the yield curve, respectively.[2] Repeating this for every period, they are left with a time series of estimated factors.[3] Each factor is then modelled and forecasted separately as an univariate AR(1) process using a sequence of expanding windows, starting with an initial sample of five years. This method relates nicely to our return forecasting method, and has also a series of advantages, notably its tractability. The dynamic Nelson-Siegel one-year yield forecast is thus the summation of each latent factor forecast times its loading:

$$\widehat{y}_{t+1}^{(n)} = \widehat{L}_{t+1} + \widehat{S}_{t+1} \left(\frac{1 - e^{-\lambda n}}{\lambda n} \right) + \widehat{C}_{t+1} \left(\frac{1 - e^{-\lambda n}}{\lambda n} - e^{-\lambda n} \right). \tag{1.8}$$

Using (1.5), it is possible to show that return forecast errors using both yield forecasting approache are simply yield forecast errors multiplied by the maturity:

$$r_{s+1} - \widehat{r}_{s+1} = (n-1) \left(y_{s+1}^{(n-1)} - \widehat{y}_{s+1}^{(n-1)} \right). \tag{1.9}$$

[1] Integrating these forward rates results in the corresponding zero-coupon yields.

[2] Diebold and Li (2006) define the level as the 10-year yield, the slope as the difference between the 10-year and 3- month yields, and the curvature as the twice the 2-year yield minus the sum of the 3-month and 10-year yields.

[3] Using our data sample, we found the relation between the estimated factors and the level, slope and curvature definition of Diebold and Li (2006) weak (considering that the dataset contains only one to five-year yields).

1.2.3 Forecasting Performance

We evaluate the performance of the forecasting exercise with an out-of sample R-square as in Goyal and Welch (2003) and Ferreira and Santa-Clara (2011).

$$R^2 = 1 - \frac{MSE_R}{MSE_M}.$$

MSE_R is the mean squared error of the out-of-sample predictions from the model:

$$MSE_R = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} \left(\mathbf{r}_{s+1} - \widehat{\mathbf{r}}_{s+1}\right),$$

and MSE_M is the mean squared error of the historical sample mean:

$$MSE_M = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} \left(\mathbf{r}_{s+1} - \overline{\mathbf{r}}_{s}\right),$$

where $\overline{\mathbf{r}}_s$ is the historical mean of excess log returns up to time s. Note that the out-of-sample R-square will take negative values when the historical sample mean predicts returns better than the model.

1.3 Empirical Analysis

1.3.1 Data and Predictors

We use the Fama-Bliss dataset (available from CRSP) of monthly one- through five-year zero coupon bond prices, from January 1965 to December 2010. Monthly values are derived from end-of-month observations. From this dataset we compute one-year holding period returns. The one-year risk-free rate is assumed to be the one-year zero-coupon yield. Table 2.1 presents summary statistics for the yield curve and the one-year holding period returns. The descriptive statistics of yields in panel A present stylised facts known to the yield curve. The yield curve is on average upward sloping, long yields are less volatile than short yields, and yields of all maturities are very persistent (long yields are more persistent than short yields). The bottom three rows of the descriptive statistics show autocorrelations at displacements of 1, 12, and 24 months. For our one-year forecast horizon, autocorrelations are around 0.80 across maturities. The descriptive statistics of bon returns in Panel B shows that an investor who purchased 5-year bonds and sold them 12 months later earned on average 1.94% yearly in excess of the risk free rate. The maximum holding period excess return was 16.89% for the 5-year bond in March 31, 1986. Holding period returns are more volatile than yields. The standard deviation of the 5-year bond yield is 2.66%, while the standard deviation of the one-year holding period return on a 5-year bond is 5.74%. This and the high persistence of yields is what makes our approach work well.

We also collect data on bond characteristics (issue date, coupon rate, maturity, callability features) as well as monthly observations of face value outstanding from CRSP, from January 1965 to December 2010. For predictive variables that are not derived bond yields, we use the dataset updated maintained by Goyal and Welch (2003).[1] From this dataset we draw monthly observations of the market return (proxied by the S&P 500 continuously compounded returns, including dividends), and inflation.

The bond return predictors are:

Forward rate (f): the log 1-year forward rate at time t for loans between time $t+n-1$ and $t+n$.

Forward spread (fs): The forward spread is defined as the difference between the forward rate and the one-year yield.

Term spread (ts): The n-year term spread is defined as the difference between the n-year bond yield and the one-year yield.

Real bond yield (ry): The real bond yield is defined as the difference between the current yield and annual year-on-year inflation:

Bond beta: The bond beta is the slope coefficient from a regression of excess bond returns on excess stock stock returns.

Inverse relative wealth (InvRelw): The inverse relative wealth is the ratio of past to current real wealth. Ilmanen (1995) motivates the use of this measure as a proxy for time-varying risk aversion. Asset risk premia should be positively related to aggregate relative risk aversion levels as suggested in Constantinides (1990), and Cochrane and Campbell (1999). We follow Ilmanen (1995) and use the real stock market index as a empirical proxy for aggregate wealth. The InvRelw is computed using an exponentially weighted average of past wealth levels, with a smoothing coefficient value of 0.90.[2]

[1] This data is maintained and updated by Amit Goyal at http://www.hec.unil.ch/agoyal/

[2] $InvRelw_t \equiv \frac{\sum_{i=1}^{t-1} 0.9^{i-1} W_{t-i}}{W_t}$.

Cochrane and Piazzesi factor ($\gamma^\intercal \mathbf{f}_t$): The Cochrane and Piazzesi factor is estimated by running a regression of the average (across maturities) excess return $\overline{r}_{t+1}$ on all forward rates:

$$\overline{r}_{t+1} = \gamma^\intercal \mathbf{f}_t + \overline{\epsilon}_{\mathbf{t+1}}. \tag{1.10}$$

The Cochrane and Piazzesi factor is defined as $\widehat{\gamma}^\intercal \mathbf{f}_t$.

Relative supply of long-term bonds ($D_t^{10^+}/D_t$): Greenwood and Vayanos (2008) define the relative supply of long- to short-term bonds as the ratio between total outstanding payments in ten years or longer and total outstanding payments.[1] For this purpose, we replicate their indicator by collecting data on every U.S. government bond that was issued from 1965 from the CRSP historical bond database.

1.3.2 Results

We first reproduce the models of the existing literature.[2] Table 1.2 presents the estimated coefficients from the Fama and Bliss (1987) regressions of bond excess returns on the forward spread, test statistics and out-of-sample R-squares. We follow Cochrane and Piazzesi (2005) and compute standard errors following Newey-West (1987), allowing for 18 months of lags. This covariance matrix is positive definite in any sample, giving more weight to more recent lags.[3] We find results very similar to those of Fama and Bliss (1987). The forward spread coefficients are close to one, and the level coefficients are close to zero. Cochrane (2005) shows that a one-for-one variation of the expected excess return on a n-maturity bond and the forward spread on the same-maturity bond is equivalent to the n-year yield following a random walk.[4] The χ^2 statistics for joint parameter significance are above the 1-percent critical value of 6.64, with the exception of the coefficient for the 5-year return, which is only significant at the 5-percent level. The in-sample R-squares are 11.6%, 13.2%, 14.9% and 7.3% for the 2, 3, 4 and 5-year bond excess returns, respectively. These results are also similar to the regressions in

[1] Outstanding payments at time t are the sum of principal and coupon payments from all bonds, bills, and notes that were issued at time t or before and have not yet retired, scaled by their face value outstanding.

[2] We use the Matlab code from Cochrane and Piazzesi (2005), available at John Cochrane's website: http://faculty.chicagobooth.edu/john.cochrane/research/.

[3] We also compute Hansen-Hodrick (1980) standard errors that explicitly control for overlapping observations, imposing equal weights on the first 12 lags. However, because both measures were very similar, we use only the first. We use the Newey-West 18 lag standard errors for joint test statistics.

[4] Replace the bond excess return definition in (1.6).

Cochrane and Piazzesi (2005) who replicate the Fama-Bliss regression to compare it to their single-factor model. Both studies argue that positive slope coefficients and high R-squares prove that returns are predictable and thus violate the Expectations Hypothesis. Figure 1.1 plots the 5-year bond excess return forecasts of the Fama and Bliss and random walk approaches against actual excess returns. It is clear from this figure that the Fama and Bliss excess return forecast converges to the random walk excess return forecast over the sample period and that the random walk approach performs much better at the beginning of the sample. Out-of-sample R-squares for the Fama and Bliss approach are 2.4%, 7.0%, 8.5% and 4.0% for the 2, 3, 4 and 5 years, respectively.

Second, we analyse the regressions with the predictive variables of Ilmanen (1995). Ilmanen proposed using variables linked to yields, such as the yield spread and real yields, but also variable linked to risk and risk aversion, such as the bond beta and the inverse relative wealth. Like our study, Ilmanen assesses the out-of-sample power of his regression. Using one-month holding period returns, he finds out-of-sample R-squares up to 12%.[1] Table 1.3 presents the results for the one-year holding period return regressions. Regression coefficients are jointly significant at the 1-percent level. Compared to Fama and Bliss, in-sample R-squares are much higher, around 30%. On the other hand, the out-of-sample R-squares produced using Ilmanen's regressions are much lower, at -5.6%, 1.3%, 2.4% -2.5% for the 2, 3, 4 and 5 years, respectively.

Table 1.4 presents results for the Cochrane and Piazzesi (2005) regressions. They use a two-step procedure in their regressions in order to capture the single-factor structure for return forecasts. They find that the this two-step procedure has little effect on the excess return regression estimated coefficients, but adds economic meaning. Their return-forecasting single tent-shaped factor is unrelated to the usual term structure factors of level, slope and curvature. The table shows the regression coefficients and test statistics for the holding period excess return regressions on the single-factor. Coefficients are jointly significant and in-sample R-squares are as high as 32.6%. The out-of-sample R-squares on the other hand are negative for all maturities. Out-of-sample R-squares are as low as -8.5%, underperforming the historical average return.

[1] The yield forecasting approach based on yields following a random walk produces out-of-samples R-squares as high as 92%, using monthly bond returns. This is because yield forecasting errors decrease with the forecasting horizon, making return forecasts more accurate.

Greenwood and Vayanos (2008) find that, in accordance to a preferred habitat theory of the term structure, the supply of long- relative to short-term bonds helps to explain future bond excess returns. In addition to this variable, they also include the Cochrane and Piazzesi (2005) single-factor and the term spread. Table 1.5 presents the results of the Greenwood and Vayanos regression.[1] All regression coefficients are jointly significant at the 1-percent level. Like the previous regression, in-sample R-squares are high, up to 30.9% for the 4-year return, but out-of-sample R-squares are negative for all maturities.

The models above do badly in predicting excess returns out of sample. With exception of the Fama and Bliss method which converges to the random walk excess return forecast, the models fall short of a forecast based on the historical sample mean. Table 1.6 presents the results for the yield forecasting approach. Panels A and B present summary statistics of yield forecasts and out-of-sample R-squares for the random walk and dynamic Nelson Siegel approaches, respectively. For both forecasting methods, mean residuals are negative. This reflects the fact that the average yield curve is upward sloping, and both methods have poor performance predicting negative slopes of the yield curve. The summary statistics results are reflected in the overall out-of-sample performance. The random walk predicts excess returns with R-squares of up to 15.3%, while the dynamic Nelson-Siegel predicts excess returns with R-squares of up to 23.6%. Figure 1.2 plots the 5-year bond excess return forecasts against the actual excess returns for the Cochrane and Piazzesi's and dynamic Nelson-Siegel approaches. From figures 1.1 and 1.2 it is possible to see that the yield forecasting approach performs much better than the traditional predictive regression in forecasting bond excess returns.

1.3.3 Robustness Check

As a robustness check of our findings, we conduct the out-of-sample exercise for the Fama and Bliss (1987) and dynamic Nelson Siegel methods on the Gurkaynak, Sack, and Wright dataset of monthly one- to twenty-year zero coupon bond prices (available from the Federal Reserve Board), from July 1982 to December 2010.[2,3] This dataset enables

[1]The regression using the term spread instead of the single-factor yielded similar results.

[2]Thirty-year zero-coupon bond yields are available only from November 1985. We incorporate the twenty-one- through thirty-year yields in the dynamic Nelson-Siegel forecast, from January 1986 onwards.

[3]Note also that the the maturities we use are equally spaced, unlike Diebold and Li (2006) which use a dataset containing maturities with lower frequencies at the short end of the yield curve. Implicitly, they are weighting the short end of the yield curve more when fitting their model. Although using

us to evaluate the models using a more representative cross-section of the term structure of bonds, albeit for a shorter time series. For simplification, we report only the 3, 5, 7, 10, 15 and 20-year returns. Despite being constructed using different bootstrapping and smoothing techniques than the Fama-Bliss dataset, one to five-year yields for the overlapping period are very close. This makes three and five-year return regressions comparable across datasets. Panel A in Table 1.7 presents the estimated coefficients from the Fama and Bliss regressions of bond excess returns on the forward spread, test statistics and out-of-sample R-squares. The Fama and Bliss model performs much worse in this second dataset. Forward spread coefficients are far from one even for the three and five-year bond excess return regression that share a common sample. Regression coefficients fail the joint significance test at any level, and in-sample R-squares are only marginally positive. The small sample bias is more acute in the second, smaller, sample. Negative out-of-sample R-squares (as low as -33%) is evidence that the models overfit in sample, and thus underperform out of sample.

Panel B presents summary statistics of yield forecasts and out-of-sample R-squares for the dynamic Nelson Siegel approach. Mean residuals are negative, reflecting the fact that the model is still deficient predicting negative slopes for the yield curve. Yield RMSE are much smaller than for the Fama-Bliss dataset. Clearly, the dynamic Nelson Siegel approach benefits from estimating L_t, S_t and C_t using a larger cross-section of bond yields and thus produces better yield forecasts. The summary statistics results are reflected in the overall out-of-sample performance. Out-of-sample R-squares are positive for all maturities, reaching values of up to 29.8% for the seven-year bond return. Interestingly, in spite of having lower yield forecasting RMSE than for other maturities, the out-of-sample R-squares for the 15- and 20-year bond returns are only 15.3% and 8.7%, respectively. This is because the excess return forecasting error for the n-maturity bond is $(n-1)$ times the yield forecasting error, and MSE are penalised by $(n-1)^2$. Thus even though the yield forecasting improves with maturity, out-of-sample R-squares will be smaller.

---

equally spaced is not necessarily optimal (we underestimate the short end variation of the yield curve), our aim is just to produce better forecasts than the random walk approach, to show the power of the yield forecasting approach for bond returns.

1.4 Conclusion

Bond returns have long been thought to be predictable. Nevertheless, the existing literature lacks out-of-sample evidence. Our paper closes this gap. We assess the predictive power for one-year holding period bond excess return regressions of Fama and Bliss (1987), Ilmanen (1995), Cochrane and Piazzesi (2005) and , Greenwood and Vayanos (2008). We find that high in-sample R-squares are a misleading indication of out-of-sample predictability. Problems arising from sampling errors in the data cause regressions to overfit in sample, and underperform out of sample.

We propose a new approach for predicting bond excess returns. Instead of forecasting returns directly, we forecast bond yields and replace them in the bond excess return definition. We use two bond yield forecasting methods: a random walk and a dynamic Nelson-Siegel approach proposed by Diebold and Li (2006). Both approaches outperform return forecasts based of the existing literature. An investor who used a simple random walk on yields would have predicted bond excess returns with R-squares of up to 15%, while a dynamic Nelson-Siegel approach would have produced R-squares up to 30%.

1.5 Tables and Figures

Table 1.1: Yield Curve Summary Statistics

This table reports summary statistics for yields and one-year holding period excess returns of Fama-Bliss yields from CRSP, from January 1965 to December 2010, with bond yields of maturities of 2 to 5 years. The last three rows of Panel A contain sample autocorrelations at displacements of 1, 12, and 24 months.

Panel A. Yield curve

	Maturities n				
	1	2	3	4	5
Mean (%)	5.965	6.183	6.365	6.517	6.617
Max (%)	15.812	15.639	15.571	15.835	15.010
Min (%)	0.248	0.314	0.513	0.844	1.179
Std. Dev	2.985	2.906	2.807	2.733	2.659
$\rho(1)$	0.981	0.984	0.984	0.985	0.986
$\rho(12)$	0.789	0.809	0.821	0.827	0.838
$\rho(24)$	0.567	0.624	0.657	0.678	0.699

Panel B. One-year holding period excess returns

	1	2	3	4	5
Mean (%)		0.500	0.880	1.151	1.209
Max (%)		5.968	10.261	14.381	16.889
Min (%)		-5.595	-10.426	-13.545	-17.548
Std. Dev		1.837	3.368	4.674	5.737

Table 1.2: Fama and Bliss Excess Return Regressions

This table reports coefficient estimates and corresponding statistics for regressions of one-year holding period excess returns on constant maturity Treasury Bonds on the forward spread $(f_t(n) - y_t^1)$. The in-sample R-squared is estimated over the full sample period. The out-of-sample R-squared in the bottom row compares the forecast error of the regression versus the forecast error of the historical mean. The sample period is from January 1965 to December 2010, and comprises bond returns of maturities of 2 to 5 years.

Fama and Bliss Excess Return Regressions

	Maturities n			
	2	3	4	5
Constant	0.001	0.001	-0.002	0.002
	(0.003)	(0.005)	(0.007)	(0.009)
$f_t^{(n)} - y_t^1$	0.831	1.130	1.367	1.032
	(0.236)	(0.310)	(0.376)	(0.433)
χ^2	12.539	13.468	13.520	6.121
$p-$val	[0.000]	[0.000]	[0.000]	[0.000]
R^2 (%)	11.6	13.2	14.9	7.3
OOS R^2 (%)	2.4	7.0	8.5	4.0

Notes : The regressions are $\mathbf{r}_t = \beta^\intercal \mathbf{X}_{t-1} + \varepsilon_t$. Standard errors are in parenthesis, probability values in brackets.The 5-percent and 1-percent critical values for a $\chi^2(1)$ are 3.84 and 6.64. All standard errors are Newey-West adjusted, with maximum lag of 18.

Table 1.3: Ilmanen Excess Return Regressions

This table reports coefficient estimates and corresponding statistics for regressions of one-year holding period excess returns on constant maturity Treasury Bonds on a set of forecasting variables at monthly frequency. These are the the Bond beta, Inverse Relative Wealth (InvRelw), the real yield ($y_t(n) - \pi$) and the term spread ($y_t(n) - y_t^1$). The in-sample R-squared is estimated over the full sample period. The out-of-sample R-squared in the bottom row compares the forecast error of the regression versus the forecast error of the historical mean. The sample period is from January 1965 to December 2010, and comprises bond returns of maturities of 2 to 5 years.

Ilmanen Excess Return Regressions

	Maturities n			
	2	3	4	5
Constant	-0.078	-0.140	-0.192	-0.231
	(0.020)	(0.038)	(0.055)	(0.070)
Bond Beta	-0.013	-0.021	-0.026	-0.025
	(0.022)	(0.021)	(0.023)	(0.025)
InvRelw	0.074	0.131	0.176	0.211
	(0.019)	(0.035)	(0.052)	(0.065)
$y_t^{(n)} - \pi$	0.359	0.673	0.947	1.180
	(0.109)	(0.199)	(0.279)	(0.358)
$y_t^{(n)} - y_t^1$	1.228	1.332	1.505	1.341
	(0.453)	(0.517)	(0.553)	(0.609)
χ^2	28.438	29.721	33.856	29.725
$p-$val	[0.000]	[0.000]	[0.000]	[0.000]
R^2 (%)	27.2	28.7	31.0	29.0
OOS R^2 (%)	-5.6	1.3	2.4	-2.5

Notes : The regressions are $\mathbf{r}_t = \beta^\intercal \mathbf{X}_{t-1} + \varepsilon_t$. Standard errors are in parenthesis, probability values in brackets.The 5-percent and 1-percent critical values for a $\chi^2(4)$ are 9.49 and 13.28, respectively. All standard errors are Newey-West adjusted, with maximum lag of 18.

Table 1.4: Cochrane and Piazzesi Excess Return Regressions

This table reports coefficient estimates and corresponding statistics for the regressions of the single-factor ($\gamma^\intercal \mathbf{f}_t$) on each bond excess return. The in-sample R-squared is estimated over the full sample period. The out-of-sample R-squared in the bottom row compares the forecast error of the regression versus the forecast error of the historical mean. The sample period is from January 1965 to December 2010, and comprises bond returns of maturities of 2 to 5 years.

Cochrane and Piazzesi Regressions

	Maturities n			
	2	3	4	5
$\gamma^\intercal \mathbf{f}_t$	0.460	0.859	1.236	1.444
	(0.028)	(0.025)	(0.026)	(0.036)
χ^2	45.351	44.182	47.596	42.574
$p-$val	[0.000]	[0.000]	[0.000]	[0.000]
R^2 (%)	26.9	29.0	32.6	30.8
OOS R^2 (%)	-8.5	-6.3	-3.1	-3.0

Notes : The regressions are $r_{t+1}^{(n)} = b_n \left(\gamma^\intercal \mathbf{f}_t\right) + \epsilon_{t+1}^{(n)}$. Standard errors are in parenthesis, probability values in brackets.The 5-percent and 1-percent critical values for a $\chi^2(1)$ are 3.84 and 6.64. All standard errors are Newey-West adjusted, with maximum lag of 18.

Table 1.5: Greenwood and Vayanos Excess Return Regressions

This table reports coefficient estimates and corresponding statistics for regressions of one-year holding period excess returns on constant maturity Treasury Bonds on a set of forecasting variables at monthly frequency. These are the the supply of long- relative to short-term bonds ($D10_t^+/D_t$) and the Cochrane-Piazzesi single-factor ($\gamma^\intercal \mathbf{f}_t$). The in-sample R-squared is estimated over the full sample period. The out-of-sample R-squared in the bottom row compares the forecast error of the regression versus the forecast error of the historical mean. The sample period is from January 1965 to December 2010, and comprises bond returns of maturities of 2 to 5 years.

Greenwood and Vayanos Excess Return Regressions

	Maturities n			
	2	3	4	5
Constant	-0.007	-0.010	-0.014	-0.016
	(0.007)	(0.012)	(0.016)	(0.019)
$D_t^{10^+}/D_t$	0.024	0.034	0.041	0.044
	(0.023)	(0.039)	(0.050)	(0.059)
$\gamma^\intercal \mathbf{f}_t$	0.418	0.815	1.217	1.465
	(0.086)	(0.167)	(0.228)	(0.283)
$\chi^2(1)$	30.432	28.027	32.255	29.802
$p-$val	[0.000]	[0.000]	[0.000]	[0.000]
R^2 (%)	26.1	27.4	30.9	29.1
OOS R^2 (%)	-11.0	-5.3	-0.6	-2.8

Notes : The regressions are $\mathbf{r}_t = \beta^\intercal \mathbf{X}_{t-1} + \varepsilon_t$. Standard errors are in parenthesis, probability values in brackets.The 5-percent and 1-percent critical values for a $\chi^2(2)$ are 5.99 and 9.21. All standard errors are Newey-West adjusted, with maximum lag of 18.

Table 1.6: Yield Forecasting Excess Return Predictions

This table reports summary statistics for the yield forecasting errors $(y_s(n-1) - \widehat{y}_{t+1}(n-1))$ and out-of-sample R-squares for the yield forecasting regressions. Panel A and B report results for the random walk and the dynamic Nelson Siegel approaches, respectively. The out-of-sample R-squares for the one-year holding period excess returns in the bottom rows compare the forecast error of each yield forecasting method versus the forecast error of the historical mean. The sample period is from January 1965 to December 2010, and comprises bond returns of maturities of 2 to 5 years.

Panel A. Random walk

	Maturities n			
	2	3	4	5
Mean (%)	-0.165	-0.162	-0.155	-0.146
Max (%)	6.403	5.636	4.696	4.639
Min (%)	-5.333	-4.649	-4.340	-4.340
RMSE (%)	1.809	1.639	1.503	1.431
OOS R^2(%)	12.2	14.0	15.3	9.5

Panel B. Dynamic Nelson Siegel

	2	3	4	5
Mean (%)	-0.152	-0.135	-0.131	-0.111
Max (%)	6.975	6.375	5.392	5.090
Min (%)	-4.142	-3.932	-3.762	-3.867
RMSE (%)	1.746	1.569	1.428	1.351
OOS R^2(%)	18.3	21.2	23.6	19.3

Table 1.7: Fama and Bliss versus Dynamic Nelson Siegel Excess Return Predictions

This table reports one-year holding period excess return predictions for the Fama and Bliss and the Dynamic Nelson Siegel approaches on the Gurkaynak, Sack, and Wright (Federal Reserve Board) sample. Panel A reports coefficient estimates and corresponding statistics for the Fama and Bliss regressions of one-year holding period excess returns on constant maturity Treasury Bonds on the forward spread $(f_t(n) - y_t^1)$. The in-sample R-squared is estimated over the full sample period. Panel B reports summary statistics for the yield forecasting errors $(y_s(n-1) - \widehat{y}_{t+1}(n-1))$ of the dynamic Nelson Siegel approach. The out-of-sample R-squares for the one-year holding period excess returns in the bottom rows compare the forecast error of each yield forecasting method versus the forecast error of the historical mean. The sample is period from July 1982 to December 2010, and comprises bond yields of maturities of 2, 4, 6, 9, 14 and 19 years and bond returns of maturities of 3, 5, 7, 10, 15 and 20 years.

Panel A. Fama and Bliss Excess Return Regressions

	Maturities n					
	3	5	7	10	15	20
Constant	0.017	0.027	0.036	0.046	0.050	0.052
	(0.007)	(0.012)	(0.016)	(0.022)	(0.021)	(0.020)
$f_t^{(n)} - y_t^1$	0.278	0.312	0.335	0.372	0.172	0.104
	(0.518)	(0.574)	(0.596)	(0.697)	(0.754)	(0.940)
χ^2	0.330	0.423	0.489	0.423	0.122	0.041
$p-$val	[0.566]	[0.515]	[0.484]	[0.516]	[0.727]	[0.839]
R^2 (%)	0.8	0.9	0.8	0.6	0.2	0.1
OOS R^2 (%)	-33.0	-29.0	-29.4	-29.9	-28.3	-26.9

Panel B. Dynamic Nelson Siegel Excess Return Regressions

	3	5	7	10	15	20
Mean (%)	-0.344	-0.332	-0.320	-0.300	-0.278	-0.254
Max (%)	2.015	1.660	1.316	1.480	1.735	1.910
Min (%)	-3.297	-2.727	-2.277	-2.225	-2.085	-1.911
RMSE (%)	1.209	0.975	0.830	0.733	0.698	0.673
OOS R^2 (%)	19.4	26.3	29.8	27.4	15.3	8.7

Notes : The regressions are $\mathbf{r}_t = \beta^\intercal \mathbf{X}_{t-1} + \varepsilon_t$. Standard errors are in parenthesis, probability values in brackets.The 5-percent and 1-percent critical values for a $\chi^2(1)$ are 3.84 and 6.64. All standard errors are Newey-West adjusted, with maximum lag of 18.

Figure 1.1: FB and Random Walk bond excess return forecasts

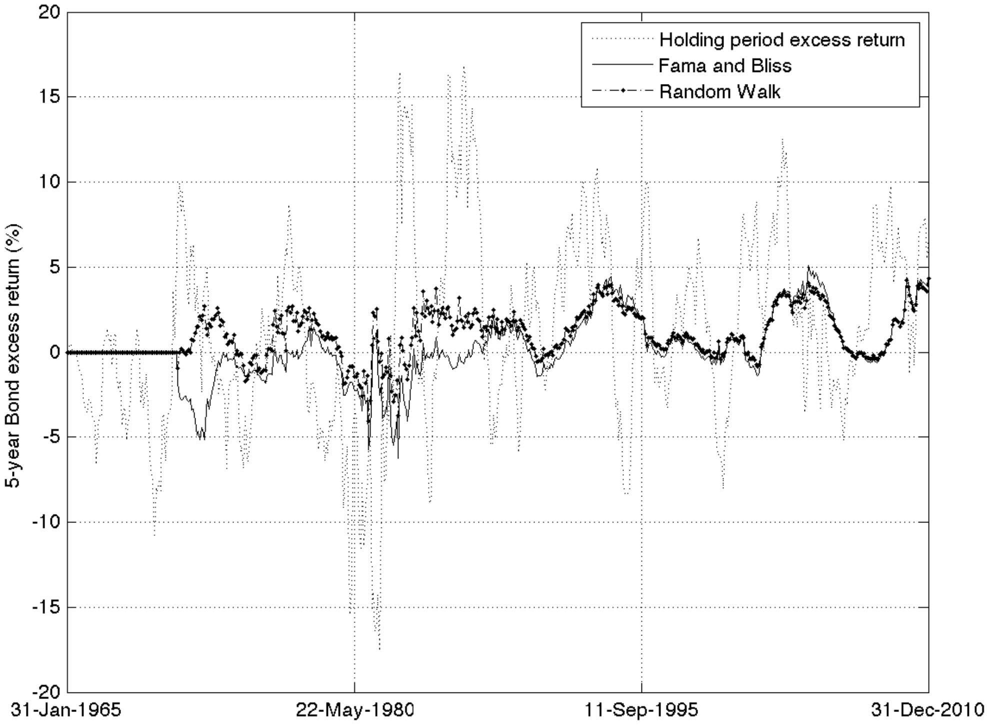
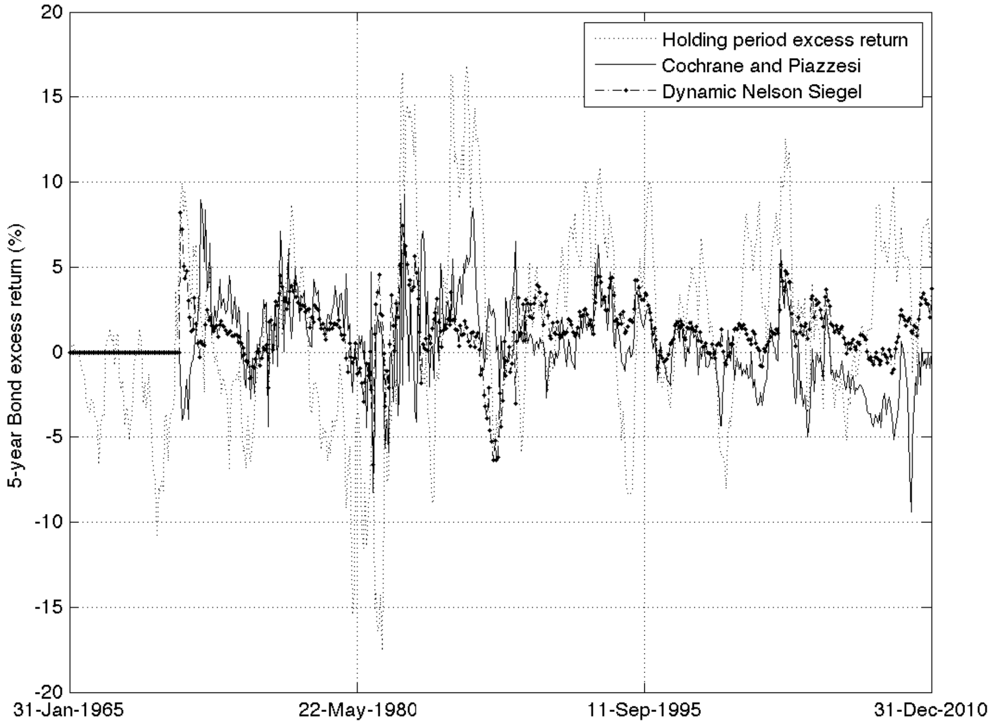


Figure 1.2: CP and DNS bond excess return forecasts



2

How You Estimate the Yield Curve Matters!

2.1 Introduction

In the past few decades a special class of term structure models termed "affine" has received a lot of attention in finance. Affine term structure (AFTS) models are based on the risk-neutral dynamics of the instantaneous short rate process. These models allow all fundamental interest rate assets (bonds and derivatives) to be priced using no-arbitrage as terms of expectations of functionals of the short rate process. Assuming no-arbitrage seems natural for bond markets since they are usually very liquid, and arbitrage opportunities are traded away immediately by investment banks. Thoroughly characterised by Duffie and Kan (1996) and Dai and Singleton (2000), this class of models encompasses the Vasicek (1977) and Cox et al. (1985a,b) seminal dynamic term structure models. It also generalises easily towards a multifactor specification of the short rate without losing its analytical tractability. Closed-form solutions for derivative prices are known for many models, adding to the desired analytical properties of this class of models.

Although these properties prove very convenient, empirical evidence against AFTS models is substantial. Backus et al. (2001) show that term premiums generated by affine models may be too low when compared to the data. Bansal and Zhou (2002) find that affine specifications are rejected by the data and propose a model that allows for regime shifts in order to account for conditional volatility and the conditional correlation across yields. Orphanides and Kim (2005) report the existence of numerous model likelihood maxima that have essentially identical fit to the data but very different implications for economic behavior. Duffee (2002) shows that AFTS models produce poor out-of-sample forecasts.[1] AFTS models have also been dismissed to price the two main interest rate derivative products: caps and swaptions. Instead, models known as "market models" are used to price these derivatives using Black's (1976) formula (Brace et al. (1997), Jamshidian (1997), Miltersen et al. (1997), Longstaff et al. (2001a,b)).[2]

In this paper, we show that the way you estimate the model matters as much as the choice of specification. We estimate a two-factor Cox et al. (1985a,b) model (CIR) on a dataset of weekly zero-coupon Euribor yields from Datastream, for the period from April 3, 2002, to October 26, 2011. This model is well known and has been extensively studied in the literature (Longstaff and Schwartz (1992), Chen and Scott (1992, 1993), Pearson and Sun (1994), Ball and Torous (1996), Duffie and Singleton (1997), Dai and Singleton (2000), Lamoureux and Witte (2002), Jagannathan et al. (2003), Duffee and Stanton (2004), Phillips and Yu (2005)). It is particularly useful because closed form expressions for the transition and marginal densities are known. This makes the model convenient to estimate using maximum likelihood and to compute derivative prices using closed form solutions. We study three basic applications of term structure models: the fitting of the yield curve, yield forecasting, and derivative pricing. For the latter, we compute cap prices using closed form solutions from Chen and Scott (1992), and then invert the cap prices and compute implied volatilities using Black's (1976) formula.[3] We then compare the implied cap volatilities from the two-factor CIR model with Euribor cap volatilities from Datastream, for the period from March 2, 2005 to October 26, 2011.

We follow an estimation method that is standard in this literature.[4] We use a state-space framework where cross-section pricing errors link observable yields to the

[1]There is only one exception. Christensen et al. (2011) develop an AFTS model based on Diebold and Li (2006). They show that the arbitrage-free restriction improves forecasts. However, little is known of this model other than its pricing and forecasting ability. Interest rate derivative pricing has not yet been developed for this model.

[2]There are only few empirical studies of AFTS models using derivative price data. Jagannathan et al. (2003) apply the CIR model for pricing caps and swaptions and find pricing errors that are too large relative to the typical bid-ask spread.

[3]The model in Chen and Scott (1992) is a special case of a two-factor CIR model analysed in Longstaff and Schwartz (1992). The advantage of Chen and Scott (1992) is that it reduces bond option expressions to univariate integrals.

[4]Ait-Sahalia and Kimmel (2010) provide a thorough four step description of the estimation procedure.

unobservable state vector of short rate factors. We maximize a joint log-likelihood that is the sum of the log-likelihood of the short rate factor dynamics under the risk-neutral probability measure and the log-likelihood of cross-section pricing errors under the physical measure. This framework makes it possible for the model to be identified under both physical ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) measures. We approximate the log-likelihood of the short rate factor dynamics using Ait-Sahalia (1999, 2008) closed-form approximations based on Hermite expansions. Additionally, we follow a market price of risk specification as in Cox et al. (1985b), which allows the drift of the state vector to be affine under both the physical and risk-neutral measures.[1]

The impact of the estimation approach in economic applications has not been studied before. We add an intermediate step before the optimisation procedure. We introduce measure-scaling weights, that sum up to one unit, in the joint log-likelihood. By varying these weights, we implicitly give more or less importance to fitting the term structure versus capturing the dynamics of interest rates. We find that these weights have great impact in the results. We show that giving more weight to $\mathbb{P}$ improves the cross-section fit and forecasting performance of the medium and long end of the Euribor yield curve. The fitting and forecasting root-mean-square errors (RMSE) for the model estimated with 90% of the weight allocated to $\mathbb{P}$ are almost double compared to those of the model estimated with only 10% of the weight allocated to $\mathbb{P}$. Forecasting RMSE on the 10-year yield are 0.4333% and 0.9568% for the model with 90% and 10% of weight allocated to $\mathbb{P}$, respectively. On the other hand, giving more weight to $\mathbb{Q}$ slightly improves pricing and forecasting performance on the short end of the Euribor yield curve, but greatly improves the pricing of cap volatilities. The 10-year cap volatility RMSE are 11.7909% and 3.0918% for the model with 10% and 70% of weight allocated to $\mathbb{Q}$, respectively. However, allocating too much weight on the hedging likelihood worsens cap volatility pricing performance. The 10-year cap volatility RMSE for the model with 90% of weight allocated on the hedging likelihood is 7.5296%.

This tradeoff is striking. A small deterioration in fitting the term structure results in a significant gain in the derivative pricing performance. This result is consistent with the

[1] This market price of risk specification is also used in most empirical studies of the CIR term structure model (Chen and Scott (1993), Pearson and Sun (1994), Lamoureux and Witte (2002), Jagannathan et al. (2003), Duffee and Stanton (2004), Phillips and Yu (2005))

results from Phillips and Yu (2005). They show that changes in CIR model parameters have little impact in bond pricing compared to pricing of European options.

Our paper proceeds as follows. In section 3.2 we describe the CIR model and the pricing, forecasting and derivative pricing applications, as well as the estimation procedure. Section 3.3 describes the data and presents the results for the model estimation using different measure-specific weights. Section 3.4 concludes.

2.2 Methodology

The central goal of this paper is to assess the performance of the two-factor CIR model, applied to Euribor rates, under different estimation approaches. We identify three main direct implementations of term structure models that give rise to numerous applications: fitting of the yield curve, yield forecasting, and derivative pricing. In practice, discount rates at exactly the desired maturities are not observed. Instead, they must be estimated from observed Libor, Swap and Futures quotes. If our model fits well a Euribor yield curve of bootstrapped rates, then it also fits well the original Euribor, Swaps and Futures quotes from which it was bootstrapped. Second, we test the forecasting performance of the model by forecasting 3-month ahead Euribor yields. Third, we test how our model prices interest rate caps of different maturities.

We begin this section by describing the CIR two-factor model for Euribor rates. We also describe how the model forecasts yields and prices caps. Lastly, we explain the estimation methodology.

2.2.1 A two-factor CIR model for the Euribor

Under the assumption of no arbitrage, the value process of a contingent claim $P(t,T)$, with terminal payoff $P(T,T)$, in the event of no default can be expressed in terms of the risk-free pricing kernel k_t as a martingale under the equivalent measure as

$$P(t,T) = E^{\mathbb{Q}}\left[e^{\int_t^T k_s ds} P(T,T)\right].$$

We assume no default. In this case, we can replace the risk-free pricing kernel k_t with the default-adjusted pricing kernel R_t.[1] Let $P(t,T)$ be the price of a bond that

[1] In our study we use Euribor rates which reflect the credit risk of lending to commercial banks in the Eurozone. Duffie and Singleton (1999) show that we can use the same models with different

pays one currency unit at maturity, without paying any intermediate coupons. R_t is the instantaneous short rate that drives the dynamics of the term structure. We assume the short rate to be the sum of two independent square root processes plus a constant,

$$R_t = r_{1t} + r_{2t} + \overline{r}.$$

The constant is added to help guarantee that interest rates are bounded away from zero.[1] The standard two-factor CIR model can be seen as a special case when $\overline{r} = 0$. The square root process under the physical measure is

$$dr_{it} = k_i(\theta_i - r_{it})dt + \sigma_i \sqrt{r_{it}} dW_{it}, \quad \text{for } i = 1, 2.$$

Where W_{it} are independent Brownian motions. It can be shown that under the risk-neutral probability measure it maintains a square root structure, with linear market prices of risk λ_i associated with each state variable (Cox et al. (1985b)),

$$dr_{it} = \overline{k}_i(\overline{\theta}_i - r_{it})dt + \sigma_i \sqrt{r_{it}} dW_{it}^{\mathbb{Q}}, \quad \overline{k}_i = k_i + \lambda_i, \quad \overline{\theta}_i = \frac{k_i \theta_i}{k_i + \lambda_i}. \tag{2.1}$$

We refer to the physical probability measure as $\mathbb{P}$, and the risk-neutral measure as $\mathbb{Q}$. The price of a discount bond is

$$P(t,T) = A_1(t,T) A_2(t,T) e^{-B_1(t,T) r_{1t} - B_2(t,T) r_{2t} - \overline{r}} \tag{2.2}$$

where

$$A_i(t,T) = \left[\frac{2 r_i e^{\left[\left(\overline{k}_i + \gamma_i\right)(T-t)\right]/2}}{\left(\overline{k}_i + \gamma_i\right)\left(e^{(T-t)\gamma_i} - 1\right) + 2\gamma_i} \right]^{\frac{2 k_i \theta_i}{\sigma_i^2}}, \tag{2.3}$$

$$B_i(t,T) = \frac{2\left(e^{(T-t)\gamma_i} - 1\right)}{\left(\overline{k}_i + \gamma_i\right)\left(e^{(T-t)u\gamma_i} - 1\right) + 2\gamma_i}, \tag{2.4}$$

and $\gamma_i = [\overline{k}_i^2 + 2\sigma_i^2]^{1/2}$. The instantaneous expected return on any default-free bond in the CIR model is

$$r_{it} + \frac{\lambda_i}{P(t,T)} \frac{\partial P(t,T)}{\partial r_{it}} = r_{it} - \lambda_i B_i(t,T) r_{it}.$$

Therefore the risk premium is positive whenever $\lambda_i < 0$.

interpretations of R_t. They argue that discounting at the adjusted short rate R_t accounts for both the probability and timing of a default event, as well as for the effect of losses on default.

[1]The short rate positivity matter has been solved in the case of the single-factor CIR model by Feller (1951). The multi-dimensional case is much less understood. Duffie and Kan (1996) and Dai and Singleton (2000) generalize vanishing conditions for multi-factor models. The usual empirical fix to this problem is to introduce a constant to the short rate (see Pearson and Sun (1994)), Duffie and Singleton (1997), Lamoureux and Witte (2002), Jagannathan et al. (2003)).

2.2.2 Interest rate forecasts

The conditional mean and variance of r_{is} conditional on r_{it} are given by

$$E[r_{it} \mid r_{is}] = r_{is}e^{-k_i(t-s)} + \theta_i\left(1 - e^{-k_i(t-s)}\right), \quad (2.5)$$

$$Var[r_{it} \mid r_{is}] = r_{is}\frac{\sigma_i^2}{k_i}\left(e^{-k_i(t-s)} - e^{-2k_i(t-s)}\right) + \theta_i\frac{\sigma_i^2}{2k_i}\left(1 - e^{-k_i(t-s)}\right)^2. \quad (2.6)$$

3-month Euribor zero-coupon yield forecasts can be computed using (2.5) and (3.3) and the definition of bond yield. We assess the forecasting performance through the root mean-squared error of Euribor yield forecasts.

2.2.3 Interest rate caps

A cap can be viewed as a payer interest rate swap contract where each payment is made only if it has positive value. The interest rate caps that we examine are written on Euribor with payments made at the end of each period and settlement periods of 3 months.

Euribor rates are rates at which deposits between banks are exchanged in the European Union interbank market. They can be seen as a simple forward rate on a defaultable bond. For the period $[T, S]$, the Euribor is defined as

$$Euribor(t, T, S) \equiv \frac{1}{S - T}\left(\frac{P(t, T)}{P(t, S)} - 1\right).$$

The forward swap rate is the rate at the fixed leg of the swap contract that makes the receiver forward swap receive zero net present value. At fixed year fractions τ (usually 3 or 6 months), the forward swap rate for the period $[\alpha, \beta]$ is

$$R_{\alpha,\beta}^{swap}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau P(t, T_i)}.$$

Cap contracts can be decomposed additively. For each period, the potential payment is the face value times $\tau[Euribor_t - R_K]^+$. The call option on the rate being capped is referred as a caplet. It is market practice is to price a caplet using Black (1976) (see Hull (2008)), which assumes a lognormal process for the Euribor. The cap contract is said to be ATM if R_K equals the forward swap rate at the relevant period.

Hull (2008) shows that a cap can be transformed into a portfolio of European puts on discount bonds. Let R_i be the value of the rate being capped. The value at time i of the payoff from the caplet that occurs at time $(i+1)$ is

$$\frac{\tau}{1+R_i}\max\left[R_i - R_K, 0\right] = (1+\tau R_K)\max\left[\frac{1}{1-\tau R_K} - \frac{1}{1-\tau R_i}, 0\right],$$

which is $1+\tau R_K$ times the payoff on a put option on a par zero-coupon bond with strike price $1/(1+\tau R_K)$. Therefore, a cap can also be considered a portfolio of put options on zero-coupon bonds. We use the second interpretation, a portfolio of put options on zero-coupon bonds, to compute the cap price for a two-factor CIR model following Chen and Scott (1992). We first compute the price of a put option on a discount bond. The integration region is given by $P(T,S) \leq K$ where $P(T,S)$ is the discount bond price. This generates a linear boundary

$$\sum_{i=1}^{2} \frac{r_{iT}}{r_i^*} \geq 1,$$

where

$$r_i^* = \frac{1}{B_i}\left(\ln\left(\prod_{i=1}^{2}\frac{A_i}{K}\right) - \bar{y}\right).$$

The price of a put option on a discount bond is given by

$$\begin{aligned} P^{put}(t,T,S,K) &= KP(r_1,r_2,t,T)\left(1-\chi^2\left(L_1,L_2,\nu_1,\nu_2,,\lambda_1^*,\lambda_2^*\right)\right) \\ &\quad -P(r_1,r_2,t,S)\left(1-\chi^2\left(L_1^*,L_2^*,\nu_1,\nu_2,,\lambda_1^0,\lambda_2^0\right)\right) \end{aligned}$$

where

$$\chi^2\left(L_1,L_2,\nu_1,\nu_2,,\lambda_1^*,\lambda_2^*\right) = \int_0^{L_2} F^*\left(L_1 - \frac{L_1}{L_2}x_2,,\nu_1,\lambda_1^*\right) f\left(x_2,\nu_2,\lambda_2^*\right) dx_2,$$

and

$$\begin{array}{cc} L_i = 2\psi_i r_i^*, & L_i^* = 2\psi_i^* r_i^*, \\ \delta_i = r_{it}\phi_i^2 \exp\left(\gamma_i\left(T-t\right)\right)/\psi_i, & \delta_i^* = r_{it}\phi_i^2 \exp\left(\gamma_i\left(T-t\right)\right)/\psi_i^*, \\ \psi_i = 2\left(\phi_i + \frac{\gamma_i + k_i^*}{\sigma_i}\right), & \psi_i = 2\left(\phi_i + \frac{\gamma_i+k_i^*}{\sigma_i} + B_i\left(T,S\right)\right), \\ \phi_i = \frac{2\gamma_i}{\sigma_i\left(\exp\left(\gamma_i\left(T-t\right)\right)-1\right)}, & \nu_i = \frac{4k_i\theta_i}{\sigma_i^2}. \end{array}$$

The χ^2 denotes a multidimensional cumulative noncentral chi-square distribution function. F and f are the distribution and density functions, respectively, of a univariate noncentral chi-square distribution. Numerical approximations to the function above can be found in Chen and Scott (1992, 1995).

2.2.4 Econometric method

We follow Ait-Sahalia and Kimmel (2010) and estimate the two-factor CIR model in four steps. These estimation steps are similar to those in Chen and Scott (1992, 1993), Duffie and Singleton (1997), Lamoureux and Witte (2002), and Jagannathan et al. (2003). First, we extract the value of the state vector R_t from a cross-section of zero-coupon yields. The state vector is not directly observable. Under the physical measure, bond prices follow the pricing equation in (3.3). It is possible to invert for the N state variables using N discount bonds at different maturities. It is usual in the literature, when using multi-factor models, to use a short and a long maturity in order to capture the different dynamics of the short end and the long end of the yield curve, and therefore better replicate the whole dynamics of the term structure. We choose the 9-month and 5-year Euribor zero-coupon yields to invert for the short rate factors. We invert for the two using the system of equations of yields,

$$\begin{pmatrix} y_1(t,\tau_{9m})\tau_{9m} \\ y_2(t,\tau_{5y})\tau_{5y} \end{pmatrix} = \begin{pmatrix} B_1(t,\tau_{9m})r_{1t} + B_2(t,\tau_{9m})r_{2t} + \overline{r} - \log A_1(t,\tau_{9m}) - \log A_2(t,\tau_{9m}) \\ B_1(t,\tau_{5y})r_{1t} + B_2(t,\tau_{5y})r_{2t} + \overline{r} - \log A_1(t,\tau_{5y}) - \log A_2(t,\tau_{5y}) \end{pmatrix}, \tag{2.7}$$

where $y_i(t,s)$ represents a zero-coupon bond yields with maturity s, assumed to be observed without error. Zero-coupon yields are affine functions of the state vector, and thus the likelihood function of yields is readily determined from the likelihood function of the state vector. Second, we compute the conditional density function for the square-root process r_i at time $t+s$, conditional on its realisation on time t,

$$f_{r_i}\left(r_{i,t}|r_{i,t-1}\right) = 2c_i f_{r_i}^{ncx2}\left(2c_i r_{i,t}; v_i, 2u_i\right) = c_i e^{-u_i - v_i}\left(\frac{v_i}{u_i}\right)^{q_i/2} I_{q_i}\left(2\left(u_i v_i\right)^{1/2}\right),$$

where f_r^{ncx2} is the conditional noncentral chi-square distribution and I_{q_i} is a modified Bessel function of the first kind and order q_i, and

$$c_i = \tfrac{2k_i}{\sigma_i^2\left(1-e^{-k_i s}\right)}, \quad u_i = c_i r_{ii,t} e^{-k_i s}, \quad v_i = c_i r i_{,t+s}\,, \quad q_i = \tfrac{2k_i\theta_i}{\sigma_i^2} - 1.$$

The joint likelihood of the short rate factors is the product of the two transition functions. Instead of using the analytical solution for the conditional density of the short rate factor in the likelihood function, we use Ait-Sahalia (1999, 2008) closed-form approximations based on Hermite expansions to the CIR likelihood function.[1] This procedure is faster and more accurate in computing the joint likelihood.

Third, we multiply this joint likelihood by a Jacobian determinant to find the likelihood of the panel of observations of the benchmark yields. As we are working with Euribor zero-coupon yields, the Jacobian J is

$$J = \frac{1}{\tau_{9m}\tau_{5y}} \begin{vmatrix} B_1(t,\tau_{9m}) & B_2(t,\tau_{9m}) \\ B_1(t,\tau_{5y}) & B_2(t,\tau_{5y}) \end{vmatrix}.$$

The log-likelihood of the short rate dynamics under the risk-neutral measure $\mathbb{Q}$ is

$$log^D = log(J^{-1}) \sum_{t=1}^{T-1} \sum_{i=1}^{2} log f_{r_i}(r_{i,t}|r_{i,t-1}). \tag{2.8}$$

We follow Chen and Scott (1993), de Jong (2000), and Duffee (2002), and assume that a second set of yields is observed with error. It is common to assume that the errors are i.i.d Normal with zero mean. The log-likelihood for measurement errors under the physical measure $\mathbb{P}$ is

$$log^C = -\frac{NT}{2} log(2\pi) - \frac{T}{2} log(\det \Sigma_t) 1 \frac{1}{2} \sum_{t=2}^{T} (\widehat{\mathbf{y}}_t - \mathbf{y}_t)' \Sigma_t^{-1} (\widehat{\mathbf{y}}_t - \mathbf{y}_t),$$

where $\mathbf{y}_t$ is a vector of observed yields and $\widehat{\mathbf{y}}_t$ is a vector of yields estimated using (3.3). Different settings can be made on these measurement errors. Either all of the yields are observed with error or only a subset of yields are observed with error. The variance terms of Σ_t is nonzero for all maturities we wish to add in the cross-section errors. In our estimation, we assume that the 6-month, 3- and 15-year Euribor zero-coupon yields are observed with error, and Σ_t is a diagonal matrix.[2]

As a fourth step we add the two log-likelihood functions to find the joint log-likelihood of the panel of all yields,

$$log L = log \mathbb{Q} + log \mathbb{P}.$$

[1] We find that these likelihood approximations produce better results than using the analytical density function. The whole algorithm and explicit expressions for the two-factor CIR likelihood approximations are described thoroughly in Ait-Sahalia (1999) and Ait-Sahalia (2008). The expressions are quite lengthy and take more than one page. Matlab codes are available at Ait-Sahalia's website at http://www.princeton.edu/ yacine/.

[2] Assuming a diagonal structure for the covariance matrix yielded better results in our estimation. The cross-covariance terms were close to zero and did not affect the results.

We estimate the model by maximising $logL$. This affine model can be seen as a state space system. The cross-section errors link observable yields to the state vector and the implied short rate factors describe the dynamics of the state vector.

This framework is necessary to identify the parameters under the risk-neutral measure $\mathbb{Q}$. Note that bond prices in (3.3) are written in terms of $\overline{k_i} = k_i + \lambda_i$, and the conditional factor density function in (3.6) are written in terms of k_i only. If we estimate the model using only $logL^{\mathbb{Q}}$, we will not able to estimate k_i and λ_i separately. The market prices of risk, λ_i identify the risk-neutral measure $\mathbb{Q}$. Therefore, if we estimate the model in this way, we will estimate the parameters under the physical measure $\mathbb{P}$. This will suffice to price and forecast bonds, since (3.3) and (2.5) are equivalent under $\mathbb{P}$ and $\mathbb{Q}$.

Conversely, if we estimate the model using only $logL^{\mathbb{P}}$ (this is equivalent to assume that all rates are observed with error), we can not invert the system of equations in (2.7) to compute the state vector. In this case we will estimate poorly the factor dynamics under the $\mathbb{Q}$. Since interest rate derivatives are priced as expectations of functionals of the process short rate under $\mathbb{Q}$, we will not be able to correctly price interest rate derivatives.

To estimate the model properly, we must use the joint log-likelihood. In other words, to estimate the model using $logL^{\mathbb{Q}}$ and $logL^{\mathbb{P}}$, the weight of each log-likelihood in $logL$ must be greater than zero. However, the magnitude of this weight has not subject of study. Previous studies usually assume that both measures enter with the same weight. The effects of changing the weights in the joint-likelihood estimation are unknown.

We introduce a measure-scaling weight alpha, in the joint log-likelihood function,

$$logL = \alpha log\mathbb{Q} + (1-\alpha)log\mathbb{P}. \tag{2.9}$$

We estimate the model using the joint log-likelihood above for different alphas in the open interval $(0,1)$. We choose alphas equal 0.1, 0.3, 0.5. 0.7, and 0.9, and assess the CIR model performance in the three basic applications for term structure models described above.

2.3 Empirical Analysis

2.3.1 Data

Euro Interbank Offered Rates (Euribor) rates are based on the average interest rates at which a panel of more than 50 European banks borrow funds from one another. There are different maturities, ranging from one week to one year. The Euribor rates are considered to be the most important reference rates in the European money market. They provide the basis for the price of Euro interest rate swaps, interest rate futures, saving accounts and mortgages. We use a dataset of Euribor weekly zero-coupon yields bootstrapped from Euribor, swaps and futures quotes, obtained from Datastream for the period from April 3, 2002, to October 26, 2011. This dataset includes Euribor zero-coupon yields with maturities ranging from 3 months to 30 years. We need only a subset of yields for our estimation. We use the 9-month and 5-year yield to invert for the short rate, and the 6-month, 3 and 15-year yields for measurement errors. In addition, we use the 3-month, 1, 2, 7, 10 and 20-year yields for our out-of-sample pricing and forecasting exercises.

Table 2.1 presents the summary statistics for the Euribor yield curve. The Euribor yield curve on average is upward sloping, long yields are less volatile than short yields, and yields of all maturities are very persistent (long yields are more persistent than short yields). The bottom three rows of the descriptive statistics show autocorrelations at displacements of 1, 3, and 12 months. For our 3-month forecast horizon, autocorrelations are around 0.80 across maturities.

We also collect weekly at-the-money cap volatilities based on the Euribor from Datastream, for the period from March 2, 2005 to October 26, 2011. Caps are said to be at the money if the strike rate equals the forward swap rate for the corresponding maturity. We choose cap volatilities with maturities of 3, 5, 7, 10, 15 and 20 years. The first two rows on Table 2.4 show the mean and standard deviations for cap volatilities.

2.3.2 Results

We estimate the model parameters for five choices of the measure weight alpha, 0.1, 0.3, 0.5, 0.7 and 0.9. A higher alpha means that the dynamics of interest rates under the risk-neutral measure $\mathbb{Q}$ have greater weight in the joint-likelihood than the pricing errors under the physical measure $\mathbb{P}$. Table 2.2 reports the parameter estimates and results for the estimation using a sample of Euribor zero-coupon yields. There is a clear pattern of results. Higher alphas yield higher hedging likelihood values, and lower pricing likelihood values. The joint log-likelihood also increases with alpha, because the hedging likelihood has greater magnitude than the pricing likelihood. It is also possible to observe a clear pattern in the model parameters. The speed of adjustment term of the first short rate factor, k_1, is of greater magnitude than k_2 and increases with alpha, which implies a faster mean reversion for r_1. Most of the variation in short-term rates is explained by the factor with higher mean reversion. This effect is the opposite for the second short rate factor. The speed of adjustment of this factor decreases with alpha. Factors with lower speed of adjustment parameters behave like a random walk and play the dominant role in the determination of long-term interest rates. The market prices of risk of the second short rate factor are negative for every value of alpha. The market prices of risk for the first short rate factor, on the other hand, change sign. They are positive for alphas up to 0.5, but become negative for higher alphas. [1] The other observable pattern is for the short rate constant added to guarantee the positivity of the short rate factors. It changes sign from negative for the two lowest values of alpha to positive and increasing with alpha. It means that the higher the alpha, the model needs to compensate more to keep short rate factors bounded away from zero.

What is the effect of the different likelihood values in the model performance? Orphanides and Kim (2005) show that AFTS models can have numerous likelihood maxima with identical fit to the data, but different economic implications. We assess the model on three economic applications: fitting of the yield curve, yield forecasting, and derivative pricing. Table 2.3, Panel A. reports root mean-squared errors (RMSE) for the two-factor CIR model on a cross-section of Euribor zero-coupon yields. The 9-month and 5-year yields that are used to invert for the short rate factors are not reported, since they have zero pricing errors. In addition, we use the 6-month, 3 and 15-year yield pricing errors to identify the model parameters under $\mathbb{Q}$. The remaining maturities are priced out of

[1] Duffee and Stanton (2004) and Phillips and Yu (2005) find estimation biases on the speed of adjustment and market price of risk parameters in one and two-factor square-root process models. Duffee and Stanton (2004) find these biases using different estimation techniques, such as the Kalman filter, maximum likelihood and efficient method of moments. These findings relate to bias in estimates of the speed of mean reversion of highly persistent processes (Ball and Torous (1996)) and of continuous time models (Chapman and Pearson (2000a), Yu and Phillips (2001)). Biases persist even when large samples are used to estimate the models.

sample. RMSE for yields of 1-year and higher maturities increases with alpha. The RMSE on the 2- and 20-year yields are 0.1344% and 0.3483%, using alpha equal to 0.1. Whereas the RMSE on the same maturities using alpha equal to 0.9 is almost double, at 0.2428% and 0.6775%. The models estimated with more weight to the $\mathbb{Q}$-likelihood have poorer performance fitting the mid and long end of the yield curve. Interestingly, model estimated with more weight to the $\mathbb{Q}$-likelihood, and thus put more weight in the short rate factor dynamics, are able to marginally price the short end of the yield curve (3- and 6-month yields) better. The RMSE on the 3-month yield is 0.2723% and 0.2331% using alpha equal to 0.1 and 0.7, respectively. The only exception is the model estimated with alpha equal to 0.9. In this setting, RMSE are higher for all yields. When alpha is too close to one, the model may be poorly estimated, as the cross-section pricing errors enter with little weight in the joint-likelihood.

With only the pricing information available above, a practitioner would be tempted to estimate the model using a lower alpha to ensure for a better cross-section fit, and then follow with forecasting and hedging. Panel B. reports RMSE for the 3-month forecasting exercise. We observe a similar picture as for the pricing errors. The models estimated with higher alphas forecast the short end of the yield curve better. RMSE on a 6-month yield decreases from 0.3238% with alpha equal to 0.1 to 0.2967% with alpha equal to 0.9. The 5-year yield, used to invert for the short rate factors, is also forecasted marginally better with higher alpha, with exception of alpha equal to 0.9. On the other hand, the models estimated with lower alphas forecast longer maturities better. This fact is in accordance with pricing errors. The models estimated with lower alphas give more weight to cross-section pricing errors, and thus fit the average shape of the yield curve better, specially at the long end of the curve. This property is then transferred to the forecasting exercise. The 10 and 20 year forecast RMSE for the model estimated with alpha equal to 0.1 is 0.4333% and 0.4484%, respectively. The RMSE for the model estimated using alpha equal to 0.9 on the same maturities is 0.9568% and 0.8028%.

The forecasting result seems to validate the practice of choosing the estimation settings that yield the best cross-section fit. We now proceed to pricing interest caps. Table 2.4 reports Euribor cap volatilities for the market and the two-factor CIR model. Caps are usually quoted in Black's volatilities, in units per 100. We price 3, 5, 7, 10, 15 and 20 year caps maturing in 6 months using the two-factor CIR model in as in Chen and Scott (1993), and then compute volatilities inverting Black's formula. This table shows that models estimated using alphas equal to 0.5 or lower underprice average cap volatilities. Underpricing decreases with alpha, and models estimated with alpha 0.7 and 0.9 slightly overprice average cap volatilities. Overall, cap volatility pricing errors are much larger than yield fitting errors. This result is in line with the results from Phillips and Yu (2005). They show that changes in a two-factor CIR model parameters have little impact in bond pricing compared to pricing of European options using this model. The model estimated using alpha equal to 0.7 has the best performance pricing caps, though both this and the model estimated using alpha equal to 0.5 produce cap RMSE that are within the 2 to 8% bid-ask volatility spread observed in the market. The average cap volatilities for the model estimated using alpha equal to 0.7 are 26.4853 and 21.1660 for the 5 and 10-year caps, whilst actual values are 24.7429 and 16.6586. The model estimated using alpha equal to 0.9 overprices cap volatilities for all maturities and has all RMSE values above 7.8%. Figures 2.1 to 2.5 show the historical fit of the model implied cap prices versus the actual prices computed using Black's formula. From these figures it is clear that the model estimated with alpha equal to 0.7 prices has the best fit of cap prices. Cap prices are close to actual values for most of the time series. Before mid 2007 the model overprices caps, increasing average cap implied volatilities. Unlike the both applications above, the interest rate derivative pricing exercise benefits from giving greater weight to the $\mathbb{Q}$-likelihood.

There is an apparent tradeoff between applications. Lower alphas give more weight to measurement errors in the joint-likelihood, and thus improve the cross-section fit of the term structure. Forecasting results depend ambiguously on both measures. Higher alphas imply that the model forecasts the maturities used in inverting for the short rate factors better. However, yields of other maturities are forecasted with greater error since the model using this setting is less able to replicate the term structure fit. Lastly, the interest rate derivative pricing exercise yields the opposite result as the term structure fit. Higher alphas imply that the parameters under the $\mathbb{Q}$ might be estimated more accurately.

Since interest rate derivative prices are computed as expectations of the short rate factors under the risk-neutral measure, this setting able to price more accurate risk neutral probabilities.

2.4 Conclusion

In this paper we assess the performance of a two-factor Cox et al. (1985a,b) model of the term structure. We estimate the model using a state-space framework, where cross-section errors link observable yields to the state vector, and the implied short rate processes describe the dynamics of the state vector. This framework is necessary to identify the model parameters under the risk-neutral measure, which in turn are necessary to accurately price interest rate derivatives. The usual way to achieve the affine term structure state-space framework in a maximum likelihood setting is to sum a log-likelihood function of pricing error to the log-likelihood of the dynamics of short rate factors.

Most studies estimate affine term structure models in this manner, choosing the estimation settings that produce the best fit of the yield curve. On a second step they use the model for different applications, such as forecasting and derivative pricing. So far, empirical results have found that derivative prices computed by short rate models produce pricing errors are large relative to the bidask spread in the market. We add an intermediary step in the estimation procedure. We give different weights to each likelihood functions and maximise the joint log-likelihood. The model performance improves depending on the choice of weights. However, it comes with a cost. The results show a strong tradeoff in the performance of economic applications for this model. Giving more weight to the physical measure likelihood improves the cross-section fit and forecasting performance. Conversely, giving more weight to the risk neutral measure likelihood improves pricing of interest rate caps. This tradeoff is asymmetric. A small deterioration of the pricing performance results in a significant gain in derivative pricing performance. We are able to price cap volatilities within the bid-ask spread bounds in the market.

2.5 Tables and Figures

Table 2.1: Euribor Zero Curve Summary Statistics

This table reports summary statistics for Euro Interbank Offered Rates zero-coupon yields obtained from Datastream, for the period from April 3, 2002 to October 26,2011. Yields have maturities 3, 6 and 9 months, and 1, 2, 3, 5, 7, 10, 15 and 20 years. The last three rows contain sample autocorrelations at displacements of 1, 3, and 12 months.

Descriptive statistics, Euribor zero curve

	3-month	6-month	9-month	1-year	2-year	3-year	5-year	7-year	10-year	15-year	20-year
Mean (%)	2.5874	2.6618	2.6589	2.7302	2.9105	3.1098	3.4489	3.7174	4.0123	4.3212	4.4419
Max (%)	5.5014	5.2411	5.2393	5.3198	5.3304	5.2763	5.1259	5.3554	5.5623	5.7716	5.8585
Min (%)	0.6239	0.9286	0.8999	1.0504	1.2392	1.3868	1.7252	2.0326	2.3631	2.6990	2.8216
Std. Dev	1.3090	1.2142	1.2106	1.1783	1.0527	0.9518	0.7987	0.7080	0.6439	0.6030	0.6243
$\rho(1)$	0.9854	0.9856	0.9821	0.9793	0.9638	0.9527	0.9375	0.9272	0.9194	0.9131	0.9156
$\rho(3)$	0.9141	0.9165	0.9020	0.8921	0.8447	0.8111	0.7688	0.7446	0.7325	0.7259	0.7380
$\rho(12)$	0.3836	0.3926	0.3963	0.3953	0.3985	0.3886	0.3601	0.3398	0.3419	0.3563	0.3928

Table 2.2: Estimation of two-factor CIR model

This table reports coefficient estimates and corresponding statistics for two-factor CIR models with different likelihood weights. All models are estimated using 6 and 9-month, 3, 5 and 15-year Euribor zero-coupon yields from Datastream, over the period from 03/04/2002 to 26/10/2011. Alpha is the likelihood weight of $logL^{\mathbb{P}}$ in $logL$. The two-factor CIR model was estimated using quasi-maximum likelihood procedure, where the likelihood of the CIR factors was approximated by a closed-form likelihood expansion as in Ait-Sahalia and Kimmel (2010). The corresponding asymptotic covariance matrix is the inverse of the Hessian matrix, which consists of the second derivatives of the joint log-likelihood function with respect to the parameters. Standard errors are in parenthesis.

		CIR parameters					$logL^{\mathbb{P}}$ parameters			Log Likelihood		
		k	θ	σ	λ	$\bar{r}$	ν_{6m}	ν_{3y}	ν_{15y}	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	$logL$
Alpha = 0.1												
	r_1	0.1819	0.1819	0.1022	0.1524	-0.0009	0.0193	0.0181	0.0359	5254.37	692.02	1148.25
		(0.0039)	(0.0006)	(0.0096)	(0.0117)	(0.0001)	(0.0027)	(0.0014)	(0.0021)			
	r_2	0.2384	0.0023	0.1414	-0.6093							
		(0.0075)	(0.0001)	(0.0013)	(0.0065)							
Alpha = 0.3												
	r_1	0.1729	0.0356	0.0912	0.1679	-0.0002	0.0175	0.0210	0.0210	5330.43	675.03	2071.65
		(0.0055)	(0.0014)	(0.0072)	(0.0130)	(0.0001)	(0.0022)	(0.0014)	(0.0021)			
	r_2	0.2288	0.0020	0.1371	-0.6162							
		(0.0129)	(0.0001)	(0.0013)	(0.0132)							
Alpha = 0.5												
	r_1	0.3523	0.0434	0.0907	0.1946	0.0020	0.0157	0.0254	0.0383	5378.17	642.11	3010.14
		(0.0195)	(0.0050)	(0.0071)	(0.0191)	(0.0002)	(0.0017)	(0.0018)	(0.0024)			
	r_2	0.1321	0.0026	0.1308	-0.5403							
		(0.0079)	(0.0001)	(0.0014)	(0.0088)							
Alpha = 0.7												
	r_1	0.4707	0.0153	0.0923	-0.1179	0.0501	0.0142	0.0358	0.0530	5464.38	499.63	3974.96
		(0.0696)	(0.0019)	(0.0081)	(0.0620)	(0.0009)	(0.0013)	(0.0058)	(0.0050)			
	r_2	0.1387	0.0010	0.1163	-0.5845							
		(0.0164)	(0.0003)	(0.0026)	(0.0361)							
Alpha = 0.9												
	r_1	0.5363	0.0144	0.1005	-0.1578	0.0723	0.0134	0.0624	0.0701	5611.14	299.060	5059.93
		(0.0546)	(0.0010)	(0.0008)	(0.0501)	(0.0001)	(0.0007)	(0.0052)	(0.0039)			
	r_2	0.0098	0.0001	0.0984	-0.5380							
		(0.0127)	(0.0016)	(0.0019)	(0.0097)							

Table 2.3: Pricing and Forecasting Errors

This table pricing and forecasting root mean-squared errors for the two-factor CIR model with different likelihood weights. All models are estimated using 6 and 9-month, 3, 5 and 15-year Euribor zero-coupon yields from Datastream, over the period from 03/04/2002 to 26/10/2011. The 9 month and 5 year yields were used to invert for the short rate factors. The forecasting window is 3 months (13 weeks).

Panel A. Pricing RMSE (%)											
	3-month	6-month	9-month	1-year	2-year	3-year	5-year	7-year	10-year	15-year	20-year
Alpha = 0.1	0.2723	0.1388	-	0.0577	0.1310	0.1344	-	0.1965	0.3110	0.1898	0.3483
Alpha = 0.3	0.2411	0.1324	-	0.0607	0.0607	0.1449	-	0.2195	0.3474	0.1909	0.3645
Alpha = 0.5	0.2150	0.1254	-	0.0655	0.1500	0.1594	-	0.2531	0.4043	0.1958	0.3935
Alpha = 0.7	0.2331	0.1194	-	0.0754	0.1743	0.1894	-	0.3379	0.5685	0.2302	0.5090
Alpha = 0.9	0.4218	0.1334	-	0.1026	0.2428	0.2696	-	0.5576	0.9650	0.3067	0.6775
Panel B. 3-month Forecasting RMSE (%)											
Alpha = 0.1	0.3653	0.3238	0.3370	0.3440	0.3941	0.3954	0.3658	0.3942	0.4333	0.3482	0.4484
Alpha = 0.3	0.3385	0.3185	0.3348	0.3425	0.3948	0.3976	0.3652	0.4045	0.4575	0.3526	0.4666
Alpha = 0.5	0.3138	0.3117	0.3315	0.3405	0.3965	0.4010	0.3641	0.3641	0.4976	0.3611	0.4983
Alpha = 0.7	0.3147	0.3026	0.3264	0.3383	0.4030	0.4114	0.3643	0.4716	0.6264	0.3991	0.6187
Alpha = 0.9	0.4578	0.2967	0.3181	0.3387	0.4331	0.4528	0.3670	0.6237	0.9568	0.4620	0.8028

Table 2.4: Cap Volatilities

This table reports actual and implied volatility statistics for the market and the two-factor CIR model. The statistics are quoted in percents.

	3-year	5-year	7-year	10-year	15-year	20-year
Market						
Mean	28.0968	24.7429	22.1764	19.7543	17.5362	16.6586
Std. Dev	13.6114	9.9808	9.9808	5.8042	4.4329	4.1521
Alpha = 0.1						
Mean	15.0979	14.7124	15.0787	13.7918	9.5113	8.9674
Std. Dev	3.8806	5.7299	5.9334	5.2556	4.1974	2.3435
RMSE	20.9967	17.8167	14.4351	11.7909	11.3380	9.7161
Alpha = 0.3						
Mean	22.1168	20.5218	18.4694	16.6536	15.0339	14.2863
Std. Dev	9.4020	5.6525	3.9631	4.0108	3.6658	3.6703
RMSE	19.0258	12.3489	9.7631	8.3070	7.2565	7.3966
Alpha = 0.5						
Mean	23.4270	20.7811	18.6189	16.5490	14.7495	13.7675
Std. Dev	9.6584	7.4294	5.6051	3.8826	2.8635	1.8511
RMSE	9.3448	6.7197	5.9492	6.4106	5.7714	5.0046
Alpha = 0.7						
Mean	29.7973	26.4853	23.7838	21.1660	18.7847	17.8731
Std. Dev	11.7545	8.8710	6.6360	4.5719	3.3147	2.2043
RMSE	5.0770	4.0640	3.5142	3.0918	3.6537	3.7861
Alpha = 0.9						
Mean	30.4084	26.7091	23.0021	20.0125	19.7318	19.9778
Std. Dev	16.6896	16.0523	13.7727	11.2714	10.0891	9.4363
RMSE	10.8483	9.4490	8.4987	7.5296	7.6920	7.8131

Figure 2.1: CIR cap prices with Alpha = 0.1

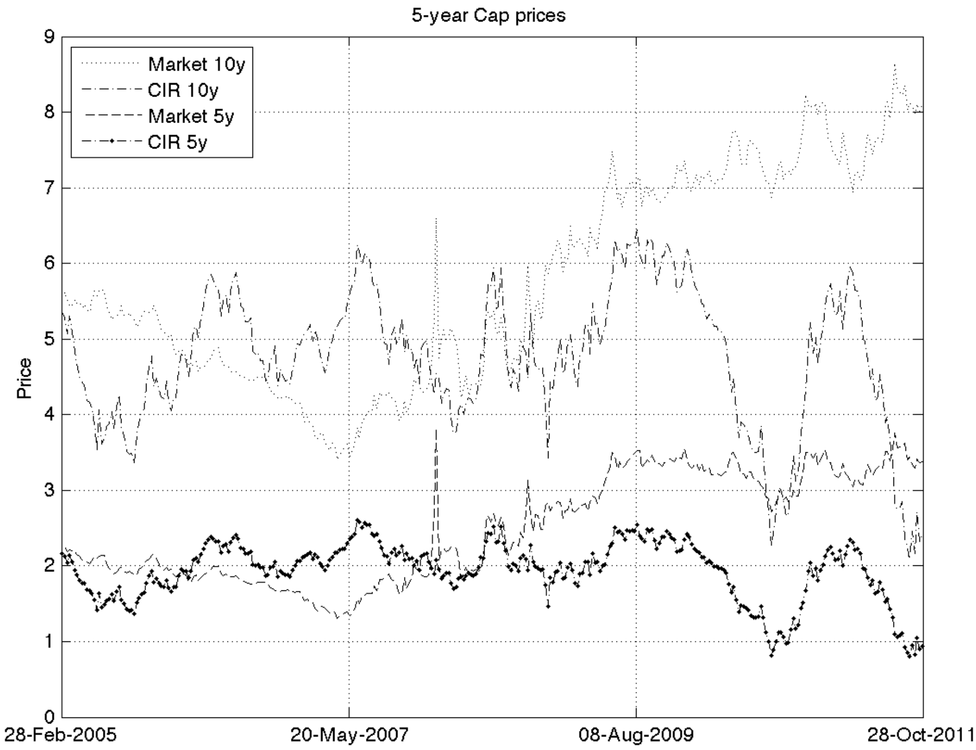


Figure 2.2: CIR cap prices with Alpha = 0.3

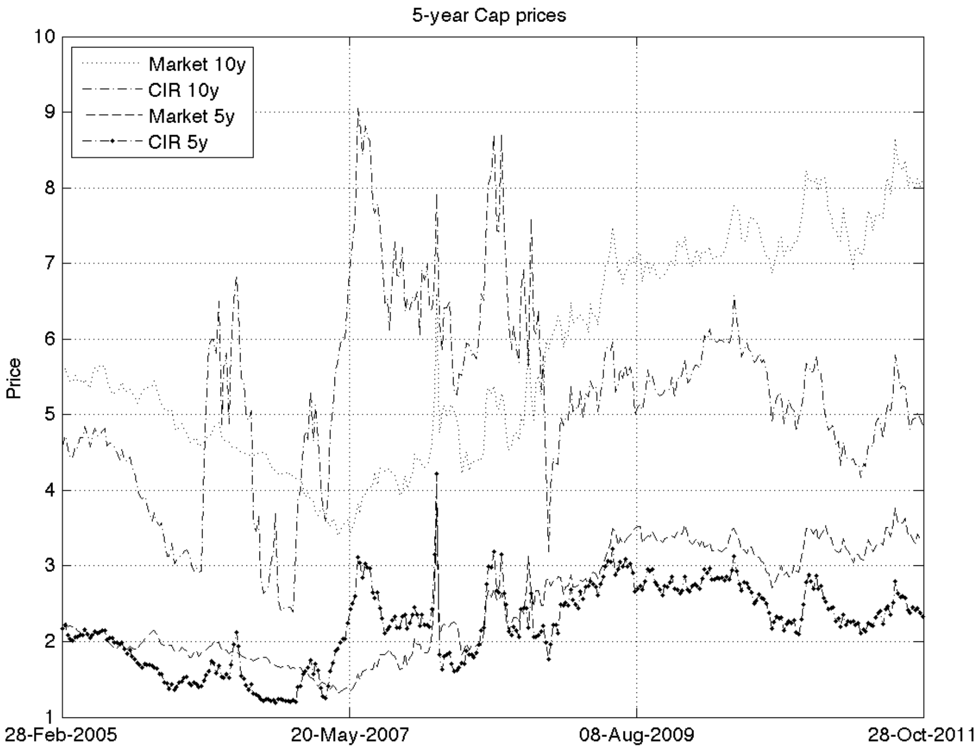


Figure 2.3: CIR cap prices with Alpha = 0.5

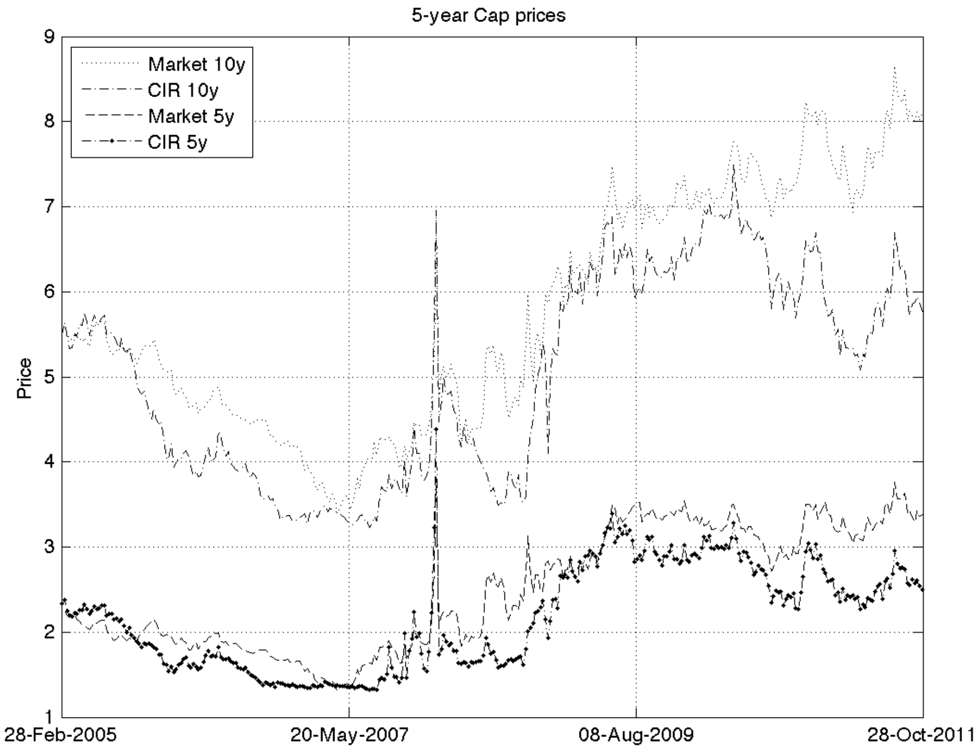


Figure 2.4: CIR cap prices with Alpha = 0.7

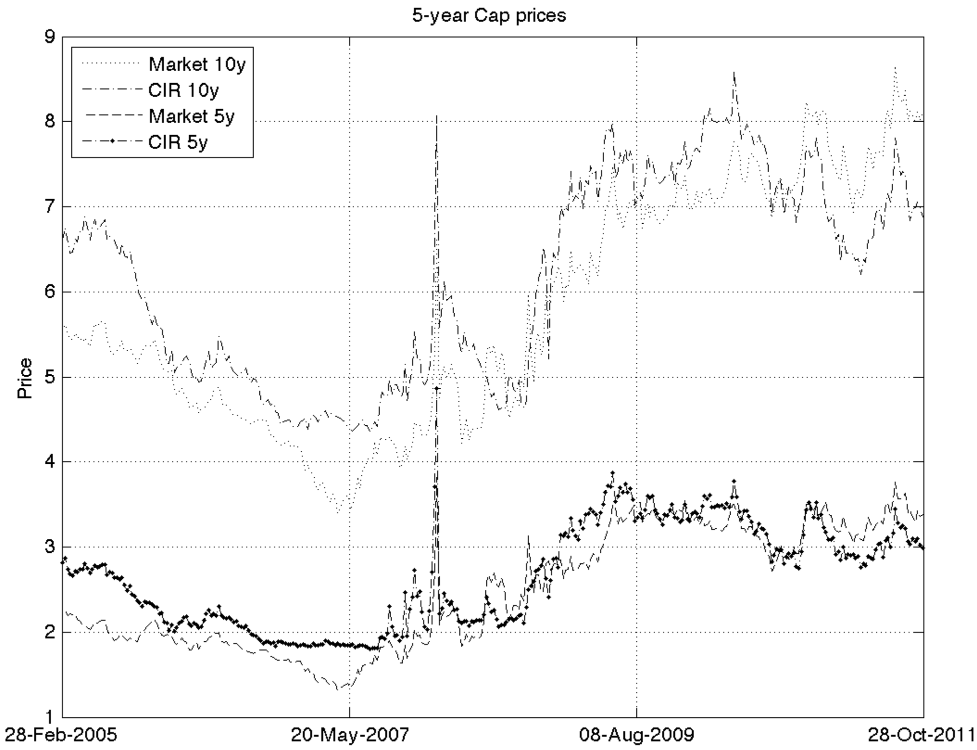
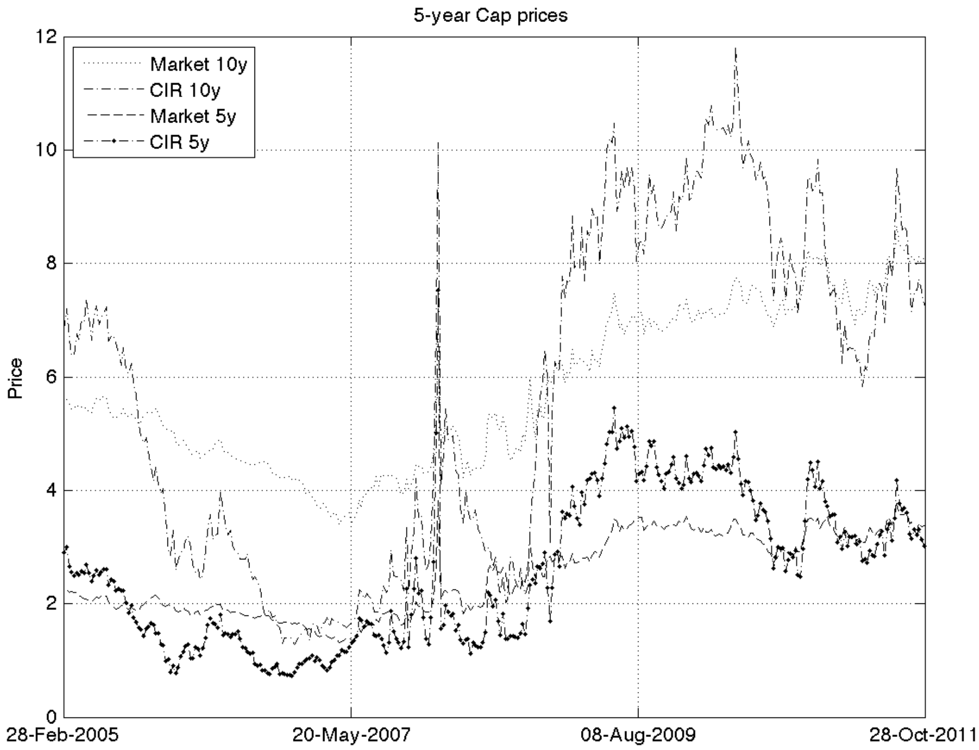


Figure 2.5: CIR cap prices with Alpha = 0.9



3

Bias and the Estimation of the CIR Term Structure Model

3.1 Introduction

Since the seminal studies of Vasicek (1977) and Cox et al. (1985a,b), many dynamic models of the term structure have been proposed. In spite of that, the Cox et al. (1985a,b) term structure model (CIR) remains the most studied term structure model in the literature. It is part of the affine class characterised by Duffie and Kan (1996), and is a special case of the Dai and Singleton (2000) family of term structure models. It generalises easily towards multiple factors and is able to generate many of the yield curve shapes observed in reality. Among its many desirable properties is tractability, as expressions of fundamental interest rate assets (bonds and derivatives) and its conditional density function are known in closed-form. For this same reason, the maximum likelihood approach is the preferred estimation technique of the CIR term structure model.[1] For a given specification of the market price of risk, which links the dynamics of the short rate factors under the risk-neutral probability measure $\mathbb{Q}$ to the dynamics under the physical probability measure $\mathbb{P}$, it is possible to maximise a joint log-likelihood function that is the sum of the log-likelihood of the short rate factor dynamics under $\mathbb{Q}$, and the log-likelihood of cross-section pricing errors under $\mathbb{P}$ to find the model parameters.[2] This state-space

[1] Duffee and Stanton (2004) show that for a simple specification of one- and two-factor versions of the Vasicek (1977) and Cox et al. (1985a,b), the maximum likelihood procedure fares best between available estimation techniques that include the Kalman filter and Efficient Method of Moments.

[2] Ait-Sahalia and Kimmel (2010) provide a thorough four-step guide to the estimation procedure.

framework is used extensively throughout the literature (Chen and Scott (1993), Pearson and Sun (1994), Duffie and Singleton (1997), de Jong (2000), Lamoureux and Witte (2002), Jagannathan et al. (2003), Duffee and Stanton (2004), Phillips and Yu (2005), Ait-Sahalia and Kimmel (2010), Fichtner and Santa-Clara (2012)).

However, the CIR model is prone to finite-sample bias. Gibbons and Ramaswamy (1993) and Pearson and Sun (1994) find that the estimated parameters of speed of adjustment are too high considering the historical behaviour of bonds.[1] Using Monte Carlo simulation, Ball and Torous (1996), Chapman and Pearson (2000b), Yu and Phillips (2001), and Phillips and Yu (2005) find that parameter estimates of the physical drift imply faster mean reversion than is implied by the true parameters. In addition, they find that the speed of adjustment coefficient is especially prone to upward bias considering the typical sample sizes of most empirical work based on Libor and swap data.[2] Phillips and Yu (2005) show that this finite-sample bias not only affects the pricing of bonds but also affects interest rate derivatives as these depend crucially on the speed of adjustment parameters of the underlying diffusions. Estimating a two-factor CIR model using Euribor rates, Fichtner and Santa-Clara (2012) find that by changing the weight allocated to each measure in the joint log-likelihood they are able to improve the derivative pricing performance of the model while worsening the fit of the yield curve. Whereas most parameters are little affected by the change in the measure-scaling weight, the speed of adjustment and market price of risk parameters in the estimation vary substantially.

Ball and Torous (1996) propose to fix the finite-sample estimation bias using weighted least squares (WLS) and GMM. Chapman and Pearson (2000b) use WLS estimation based on a simple first-order discretization of the data. Nonetheless, there is little or no evidence of bias reduction. Phillips and Yu (2005) propose a bias reduction method based on the jackknife. They show how the method can be applied to the one- and two-factor

[1]Using Treasury bill data only, Gibbons and Ramaswamy (1993) estimate the speed of adjustment parameter to be 12.43, while Pearson and Sun (1994) estimate is 9.24. Both of these estimates imply a mean half-life for the estimated interest rate process (that is, the expected time for the process to return halfway to its long-term mean) of less than one month.

[2]These studies find that the finite-sample bias persists even when the sample size is quite large. Contrary to these findings, Duffee and Stanton (2004) argue that the biases disappear by incorporating cross-section errors in the estimation and using large sample sizes (1000 weeks). However, most empirical studies of affine term structure models with implications to interest rate derivative prices find that large samples Libor and swap data are not available.

CIR model parameters, and to the option prices directly. Using Monte Carlo simulation, they show that the performance of the maximum likelihood estimation improves using the jackknife. They find that the estimated speed of adjustment parameters and option prices are no longer strongly biased, although standard deviations and root-mean-square errors (RMSE) remain around the same magnitude as the true values, and higher than when the maximum likelihood is used to estimate the model.

This paper investigates the finite-sample properties of the time series estimation of the CIR model of the term structure. I follow an estimation method that is standard in affine term structure literature.[1] I use a state-space framework in which cross-section pricing errors link observable yields to the unobservable state vector of short rate factors. Second, I maximize a joint log-likelihood that is the sum of the log-likelihood of the short rate factor dynamics under the risk-neutral probability measure and the log-likelihood of cross-section pricing errors under the physical measure. This framework makes it possible for the model to be identified under both physical and risk-neutral measures. I use Monte Carlo simulation to determine the behaviour of the estimators when introducing measure-scaling weights in the joint log-likelihood function as in Fichtner and Santa-Clara (2012).[2] The simulated data sample size is five hundred weeks, which is typical for most empirical work based on Libor and swap data. I also assess the performance of the maximum likelihood estimation when noise is introduced to simulated bond yields. This represents a more realistic scenario, since discount bonds at exactly the desired maturities are not observed in practice. Instead, they must be estimated from observed bond quotes, and thus likely to contain noise. The noise is assumed to be i.i.d. across yields for all maturities.[3] I estimate the CIR model using standard deviations of bond yield noise of 5 and 10 basis points. Lastly, I study the effects of changing measure-scaling weights and the addition of i.i.d. noise to fitting of the yield curve and derivative pricing.

The results show that the typical absolute values of the likelihood functions and their partial derivatives at the true parameter values are disproportionate. In the estimation of the one-factor model conducted using bond yields without the addition of i.i.d. noise,

[1] Ait-Sahalia and Kimmel (2010) provide a thorough four step description of the estimation procedure.

[2] The weights sum up to one, and a weight of 0.5 to each likelihood is equivalent to the traditional estimation specification.

[3] The introduction of noise is similar to Duffee and Stanton (2004). However, Duffee and Stanton add noise only to bond yields assumed to be measured with error.

the value of the log-likelihood function of the short rate factor dynamics is seven times larger than the log-likelihood function of cross-section pricing errors. The partial derivatives of the log-likelihood function of the short rate factor dynamics with respect to the model parameters at the true value vary significantly, being large and positive for the speed of adjustment parameter. Because the model is estimated using bond yields without the addition of i.i.d. noise, the partial derivatives of the log-likelihood function of cross-section pricing errors with respect to the model parameters are zero. In this setting, decreasing the finite-sample bias of the CIR model becomes trivial. Imposing a weight, alpha, of 0.1 to the log-likelihood function of the short rate factor dynamics and 0.9 to the log-likelihood function of cross-section pricing errors significantly improves the performance of the maximum likelihood estimation of the one- and two-factor CIR models, as well as the yield fitting and bond option performance. The speed of adjustment and market price of risk parameters fall to around 5% of the parameter value for the one- and two-factor models, compared to approximately 15% of the model estimated using the standard setting of alpha equal to 0.5. The pricing RMSE of the bond option falls to 3% of the option value for the one-factor model, compared to approximately 15% in the standard setting. For the two-factor model, the bias reduction for the speed of adjustment and market price of risk parameters are even more drastic.

When the model is estimated using simulated bond yields added with i.i.d noise, the one-factor model performance still increases giving more weight to the log-likelihood function of cross-section pricing errors under the physical measure. The only apparent consequence is an increase in standard deviations of estimated parameters, specially the speed of adjustment and market price of risk parameters. The standard deviation for the speed of adjustment parameter quadruples from a setting without yield noise to a setting where yield noise has standard deviation of 10 basis points. For the two-factor model, however, we observe the same tradeoff as in Fichtner and Santa-Clara (2012). Giving more weight to the log-likelihood of pricing errors under $\mathbb{P}$ is no longer optimal. As noise increases, the information on the cross-section of the yield curve becomes less precise about the parameters under $\mathbb{Q}$. By giving more importance to the dynamics of the state vector, the model is capable to discount the imprecise information contained in pricing errors and the dynamics affected by noise, and estimate the true underlying parameters more accurately. Hence, the option pricing performance improves. For the model estimated with yield noise standard deviation of 5 basis points, options are best priced using the parameters estimated using alpha equal to 0.3. The bond option RMSE for the model estimated using alpha equal to 0.3 is 60% of the RMSE obtained using alpha equal to 0.1, and 50% of the RMSE obtained using alpha equal to 0.5. For the model estimated with yield noise standard deviation of 10 basis points, the optimal option pricing performance is reached using alpha equal to 0.5. The caveat is that the performance of fitting the yield curve in the cross-section always worsens when more importance is given to the dynamics. That is the tradeoff observed Fichtner and Santa-Clara (2012). The optimal choice of the measure-scaling weight depends on the application, but also depends on the data.

The findings of this paper are twofold. The bias found in Ball and Torous (1996) and Phillips and Yu (2005) is significantly reduced by giving more weight to the pricing errors in the estimation. However, this will be of little help for a practitioner aiming to price interest rate derivatives through the two-factor CIR model using market data. When yields differs from the underlying parameter model-implied yields, the practitioner faces a choice. Varying the measure-scaling weight to give more importance to the dynamics of the short rate factors in the estimation improves the bond option pricing performance. At the same time, the performance fitting the yield curve deteriorates.

This paper proceeds as follows. In the next section I describe the CIR model and the simulation and estimation procedures. Section 3.3 presents the results for the model estimation using different measure-specific weights and different values for yield noise. Section 3.4 concludes.

3.2 Methodology

Let $P(t,T)$ be the price of a bond that pays one currency unit at maturity, without paying intermediate coupons. R_t is the instantaneous short rate that drives the dynamics of the term structure. The short rate is assumed to be the sum of N independent square root processes,

$$R_t = \sum_{i=1}^{N} r_{it}.$$

The square root process r_{it} under the physical measure is

$$dr_{it} = k_i(\theta_i - r_{it})dt + \sigma_i\sqrt{r_{it}}dW_{it}. \tag{3.1}$$

Where W_{it} are independent Brownian motions. It can be shown that under the risk-neutral probability measure it maintains a square root structure, with linear market prices of risk λ_i associated with each state variable (Cox et al. (1985b) and Dai and Singleton (2000)),

$$dr_{it} = \overline{k}_i(\overline{\theta}_i - r_{it})dt + \sigma_i\sqrt{r_{it}}dW_{it}^{\mathbb{Q}}, \quad \overline{k}_i = k_i + \lambda_i, \quad \overline{\theta}_i = \frac{k_i\theta_i}{k_i + \lambda_i}. \tag{3.2}$$

I refer to the physical probability measure as $\mathbb{P}$, and the risk-neutral measure as $\mathbb{Q}$. The price of a discount bond is

$$P(t,T) = \prod_{i=1}^{N} A_i(t,T)e^{-B_i(t,T)r_{it}} \tag{3.3}$$

where

$$A_i(t,T) = \left[\frac{2r_i e^{\left[\left(\overline{k}_i+\gamma_i\right)(T-t)\right]/2}}{\left(\overline{k}_i + \gamma_i\right)\left(e^{(T-t)\gamma_i} - 1\right) + 2\gamma_i}\right]^{\frac{2k_i\theta_i}{\sigma_i^2}}, \tag{3.4}$$

$$B_i(t,T) = \frac{2\left(e^{(T-t)\gamma_i} - 1\right)}{\left(\overline{k}_i + \gamma_i\right)\left(e^{(T-t)u\gamma_i} - 1\right) + 2\gamma_i}, \tag{3.5}$$

and $\gamma_i = [\overline{k}_i^2 + 2\sigma_i^2]^{1/2}$. The conditional density function for r_i at time $t+s$, conditional on its realisation on time t, is

$$f_{r_i}\left(r_{i,t}|r_{i,t-1}\right) = 2c_i f_{r_i}^{ncx2}\left(2c_i r_{i,t}; v_i, 2u_i\right) = c_i e^{-u_i - v_i}\left(\frac{v_i}{u_i}\right)^{q_i/2} I_{q_i}\left(2\left(u_i v_i\right)^{1/2}\right), \tag{3.6}$$

where f_r^{ncx2} is the conditional noncentral chi-square distribution and I_{q_i} is a modified Bessel function of the first kind and order q_i, and

$$c_i = \frac{2k_i}{\sigma_i^2\left(1-e^{-k_i s}\right)}, \quad u_i = c_i r_{ii,t} e^{-k_i s}, \quad v_i = c_i r i_{,t+s}, \quad q_i = \frac{2k_i\theta_i}{\sigma_i^2} - 1.$$

Option prices for the discount bonds based on the Cox et al. (1985a,b) model have analytical solutions for the one- and two-factor cases. Let $C(R,t;T,s,K)$ be the price of a European call option on a discount bond with the option expiration at time T, bond maturity at time s, $s > T$, principal L and strike price K. Cox et al. (1985b) provide a solution for $C(R,t;T,s,K)$ for the one-factor version, while Chen and Scott (1992) provide a solution for $C(R,t;T,s,K)$ when R is a two-factor state vector. Chen and Scott show that although $C(R,t;T,s,K)$ does not have an analytic expression, the computation is reduced to univariate numerical integrations of cumulative density functions and probability density functions of noncentral chi-square variates. I use the same Matlab code for computing bond option prices based on the one- and two-factor CIR model as Phillips and Yu (2005).[1] Figure 3.1 replicates Figure 2 in Phillips and Yu (2005). It depicts the relation of the speed of adjustment parameter and bond and option prices. It is clear from this that the bond prices are not nearly as sensitive to changes in the speed of adjustment parameter as option prices.

I follow Ait-Sahalia and Kimmel (2010) and estimate the two-factor CIR model in four steps, with the addition of the measure-scaling weight from Fichtner and Santa-Clara (2012).[2] First, I extract the value of the state vector R_t from a cross-section of zero-coupon yields. The state vector is not directly observable. Under the physical measure, bond prices follow the pricing equation in (3.3). It is possible to invert for the N state variables using N discount bonds at different maturities. For the one-factor model, I use the two-year bond to invert for the short rate factor. For the two-factor model, I use the two and ten-year bonds to invert for the short rate factors.

Second, I compute the conditional density function for short rate as in (3.6) and multiply it by a Jacobian to find the likelihood of the panel of observations of the benchmark yields to form log-likelihood of the short rate dynamics under the risk-neutral measure $\mathbb{Q}$, $logL^{\mathbb{Q}}$.[3] Third, I follow Chen and Scott (1993) and assume that a second set of yields is observed with error. The errors are i.i.d Normal with zero mean. I compute a log-likelihood for measurement errors under the physical measure $\mathbb{P}$. For the one-factor model, I assume that the 5-year yield is observed with error. For the two-factor model, I assume that the 1 and 7-year yields are observed with error. Σ_t is a diagonal matrix.

As a fourth step I add the two log-likelihood functions to find the joint log-likelihood of the panel of all yields, introducing the measure-scaling weight alpha as in Fichtner and Santa-Clara (2012):

$$logL = \alpha log\mathbb{Q} + (1-\alpha)log\mathbb{P}. \tag{3.7}$$

I estimate the model by maximising $logL$. The joint-likelihood function can be seen as a state space system. The cross-section errors link observable yields to the state

[1] The Matlab code is provided in Jun Yu's website at http://www.mysmu.edu/faculty/yujun/research.

[2] These estimation steps for the CIR model are similar to those in Chen and Scott (1992 and 1993), Lamoureux and Witte (2002), Jagannathan et al. (2003), and Fichtner and Santa-Clara (2012).

[3] Details and equations are thoroughly explained in Jagannathan et al. (2003) and Fichtner and Santa-Clara (2012).

vector and the implied short rate factors describe the dynamics of the state vector. This framework is necessary to identify the parameters under the risk-neutral measure $\mathbb{Q}$. I estimate the model using the joint log-likelihood above for different alphas in the open interval $(0, 1)$. I choose alphas equal 0.1, 0.3, 0.5. 0.7, and 0.9, and assess the CIR model performance in fitting the yield curve and pricing bond options.

Using Monte Carlo approach I simulate one-thousand paths of the square-root diffusion dynamics in (3.1).[1] The dynamics of the short rate factors under the physical measure $\mathbb{P}$ are determined by the market price of risk. I follow a market price of risk specification as in Cox et al. (1985b) which allow the drift of the state vector to be affine under both the physical and risk-neutral measures.[2] Bond yields for maturities from one to twenty years are computed using (3.3). The sample size is five hundred weekly observations and is similar to those used in most empirical work based on Libor and swap data, which is more relevant when computing interest rate derivatives (Duffie and Singleton (1997), Jagannathan et al. (2003), and Phillips and Yu (2005), Fichtner and Santa-Clara (2012)). The parameter values for the one-factor model are taken from Phillips and Yu (2005), while the parameter values for the two-factor model are close to those found in the literature.[3]

At last, I add normally distributed noise with standard deviation $\sqrt{V}$ to simulated bond yields. Unlike Duffee and Stanton (2004) which add noise only to the yields assumed to be observed with error, I add noise across yields for all maturities. This is a more realistic scenario. In practice, discount bonds at exactly the desired maturities are not observed. Instead, they must be estimated from observed bond quotes, and thus likely to contain some noise. I assume the noise to be i.i.d. across yields for all maturities, and across time and maturity. I estimate three scenarios. First, the simulated yields are the same as the model-implied yields (no noise). Second, the standard deviation of bond yield noise is set to 5 basis points. Third, I estimate the models using standard

[1] Matlab codes are available at Ait-Sahalia's website at http://www.princeton.edu/ yacine/

[2] This market price of risk specification is also used in most empirical studies of the CIR term structure model (Chen and Scott (1993), Pearson and Sun (1994), Duffie and Singleton (1997), Lamoureux and Witte (2002), and Jagannathan et al. (2003)). Phillips and Yu (2005) use it in a few examples only.

[3] Phillips and Yu (2005) also estimates the two-factor model, but with market prices of risk set to zero. The speed of adjustment parameters they use are too small to account for non-zero market prices of risk found in the literature and produce reasonable shapes of the yield curve.

deviation of bond yield noise equal to 10 basis points. We proceed with the results in the section below.

3.3 Results

The one-factor CIR model is estimated first. The two-year bond yield is used to invert for the short rate, while the five-year bond yield is assumed to be measured with error. I choose five values for the measure-scaling weight alpha: 0.1, 0.3, 0.5, 0.7 and 0.9. A higher alpha means that the dynamics of interest rates under the risk-neutral measure $\mathbb{Q}$ has greater weight in the joint-likelihood than pricing errors under the physical measure $\mathbb{P}$. That is, one implicitly gives more importance to capturing the dynamics of interest rates versus fitting the term structure. The underlying bond yields come from 1000 simulated bond yield scenarios using the one-factor CIR model. Simulations contain 500 weeks, with maturity of one to twenty years. The true values for k, θ, σ and λ are 0.5, 0.08, 0.04, and -0.1, respectively. The true value of a one-year European call option on a three-year discount bond with a face value of \$100, strike price of \$82, and initial interest rate of 5% is 3.5411. Table 3.1 reports mean, median, standard deviation and root mean squared error (RMSE) for parameter estimates, log-likelihood values for the physical and risk-neutral measures, as well as bond option values. In this table it is possible to notice that estimations using alpha equal to 0.1 present very small deviations from true values for estimated parameters and bond option prices. For similar parameter choices, Phillips and Yu (2005) find that parameter and bond option estimates for its maximum likelihood and jackknifing approaches have RMSE, standard deviation of the same magnitude as true values. The value of the log-likelihood of the short term factor under the risk-neutral measure, $logL^{\mathbb{Q}}$ is roughly six times higher than the value the log-likelihood of pricing errors, $logL^{\mathbb{P}}$. Among all settings, the estimation of θ and σ present the smallest standard deviations and RMSE. This goes into according with Duffee and Stanton (2004) and Phillips and Yu (2005) who find very little variation from true values for these parameters. On the other hand, it is possible to observe a pattern for the the speed of adjustment and price of risk parameters. The speed of adjustment increases with alpha, while price of risk parameters decrease with alpha by roughly the same amount. Under the physical measure $\mathbb{Q}$, this pattern cancels out as $\overline{k} = k + \lambda$. In the estimation using alpha equal to 0.9, the mean of the speed of adjustment and market price of risk are 0.6052 and -0.1854, respectively, resulting in $\hat{\overline{k}} = 0.4198$, while $\overline{k} = 0.4$. Log-likelihood values and option price also present a clear pattern. $logL^{\mathbb{Q}}$ increases with alpha, while $logL^{\mathbb{P}}$ decreases. In the estimation using alpha equal to 0.9, $logL^{\mathbb{Q}}$ is roughly twenty times higher than $logL^{\mathbb{P}}$. Because k increases on average with alpha, option prices, which depend only on the parameters under $\mathbb{Q}$, decrease on average. In the estimation using alpha equal to 0.1, the mean option price is 3.5230 and RMSE is 0.1172. When alpha increases to 0.9, the mean option price falls to around \$1 below the true value, while the RMSE increases to 1.9143.

Table 3.4 presents the RMSE of bond yields for the one-factor CIR model estimation. Panel A presents the RMSE for the model estimated without noise. Like Fichtner and Santa-Clara (2012), RMSE of bond yields increases with alpha. RMSE of a ten-year bond yield is only 0.0020 for a model estimated using alpha equal to 0.1, while 0.3508 when using alpha equal to 0.9. Giving more weight to the log-likelihood of pricing errors under the physical probability measure is clearly better than estimating the model giving equal weights to each measure. Tables 3.2 and 3.3 present results for the one-factor CIR model estimated using simulated bond yields added with i.i.d noise with standard deviations equal to 5 and 10 basis points. A similar pattern of results emerges from these tables. The estimations using alpha equal to 0.1 still perform better than for other values, albeit significantly higher standard deviations and RMSE compared to the previous setting. Table 3.2 presents an increasing pattern for the speed of adjustment parameter, both for the mean and median. In Table 3.3 the median speed of adjustment parameter falls below the true value for alphas equal to 0.3, 0.5 and 0.7. As a result, standard deviations and RMSE for this setting are much higher.

Table 3.5 reports mean, median and standard deviation of partial derivatives of the one-factor CIR model likelihood functions at the true value of the coefficients, and at different values for the standard different deviation of the yield noise, $\sqrt{V}$. For the model estimated from bond yields without noise, the partial derivatives of $log\mathbb{P}$ are zero, while the partial derivatives of $logL^{\mathbb{Q}}$ at the true value of the coefficients vary significantly. It is evident from this table that while partial derivatives of the likelihood function of $log\mathbb{Q}$ are small in absolute value with respect to θ, σ and λ (with standard deviations exceeding absolute values their mean and median), the partial derivative of $log\mathbb{Q}$ with respect to k is large and positive. Giving more weight to the log-likelihood of the short term factor under the risk-neutral measure in the estimation process implicitly gives more weight to the increasing tendency of $logL^{\mathbb{Q}}$ with respect to k. It helps to explain why allocating more weight to the risk-neutral measure in the estimation procedure results in an overestimation of k. A lower alpha has the opposite effect, and the parameters following from the optimisation procedure are closer to true values. Figure 3.2 illustrates the relation of the speed of adjustment parameter and the log-likelihood functions for the one-factor CIR model for one sample of 500 weekly observations and alpha equal to 0.5. It is possible to observe that $logL^{\mathbb{Q}}$ increases with k, while $log\mathbb{P}$ is concave around the true value of k.

Second, I estimate the two-factor CIR model. The true values of k_1, θ_1, σ_1, λ_1, k_2, θ_2, σ_2 and λ_1 are 0.5, 0.07, 0.04, -0.1, 0.05, 0.02, 0.04 and -0.02, respectively. The true value of a one-year European call option on a three-year discount bond with a face value of \$100, strike price of \$82, and initial interest rate of 5% is 2.7690. Table 3.6 reports the estimation of the two-factor CIR model when no noise is added to simulated yields. Like the one-factor model, this setting results in an upward bias for the speed of adjustment parameter and a downward bias for market prices of risk that worsen with an increase in alpha, while remaining parameters do not present significant deviations from true values. As a consequence, mean and median option price values decrease. The models estimated using alpha equal to 0.1 have the best performance. The mean and median of the speed of adjustment parameter $k1$ are 0.5146 and 0.5116, while 0.0523 and 0.0516 for $k2$. The RMSE of the option price is 0.2328, compared to 1.7941 using alpha equal to 0.9. For the model estimated with alpha equal to 0.5, the value of the log-likelihood of the short term factor under the risk-neutral measure, $logL^{\mathbb{Q}}$ is roughly fifteen times higher than the value the log-likelihood of pricing errors, $logL^{\mathbb{P}}$. Table 3.9, Panel A presents RMSE for bond yields. Like previous results, the yield fitting performance is inversely proportional to alpha.

Tables 3.7 and 3.8 present results for the one-factor model estimated using simulated bond yields added with i.i.d noise with standard deviations equal to 5 and 10 basis points. Contrary to the one-factor model, the estimations made using alpha equal to 0.1 no longer have the best performance pricing bond options. The speed of adjustment parameters downward biased, and option prices present an upward bias. For the simulated bond yields added with i.i.d noise with standard deviations equal to 5 basis points, the mean bond option is 2.9605, while RMSE is 0.8869. Using alpha equal to 0.3, these values fall to 2.7891 and 0.5271, respectively, while the mean and median values of the estimated speed of adjustment parameters are 0.4979 and 0.5001. Further increasing alpha increases the upward bias in the speed of adjustment parameters, and thus decreases mean bond option values. Bond options for the model estimated using alpha equal to 0.9 have mean of 2.0655 and RMSE of 2.2615, while the mean and median speed of adjustment parameters are 0.6439 and 0.6298. The yield pricing performance, on the other hand, presents the same pattern as the one-factor model of for the two-factor model estimated using yields without noise. Table 3.9, Panel B presents yield RMSE for this estimation. It is increasing with alpha, and the RMSE on a twenty-year yield is more than double for the model estimated using alpha equal to 0.9 than for the model estimated using alpha equal to 0.1.

Table 3.8 presents similar results. The model is estimated using simulated bond yields added with i.i.d noise with standard deviations equal to 10 basis points. Estimations using alpha equal to 0.1 and 0.3 present a downward bias for the speed of adjustment parameters, and a small upward bias for market prices of risk. Although k_2's mean and median are slightly below the true value, k_1's mean and median are 0.5027 and 0.4998 for the models estimated using alpha equal to 0.5. The option pricing performance is also improved. The mean option price is 2.7510, and RMSE is 1.1403. In contrast, RMSE for the models estimated using alpha equal to 0.1 and 0.9 are 2.0407 and 1.6022, respectively. Yield RMSE are presented in Table 3.9, Panel C. and presents the same pattern as the other panels. The results above are striking. Giving more weight to the log-likelihood of pricing errors under $\mathbb{P}$ is no longer optimal, and the Figure 3.2 no longer represents the reality when the bond yield data is no longer model implied. As noise increases, the information on the cross-section of the yield curve becomes less precise about the parameters under $\mathbb{Q}$. Giving more importance to the dynamics of the state vector, the model is capable to discount the imprecise information contained in pricing errors and the dynamics affected by noise, and estimate the true underlying parameters more accurately. The caveat is that the performance of fitting the yield curve in the cross-section worsens when more importance is given to the dynamics. This is the tradeoff found in Fichtner and Santa-Clara (2012).

Table 3.10 reports mean, median and standard deviation of partial derivatives of the one-factor CIR model likelihood functions at the true value of the coefficients, and at different values for the standard different deviation of the yield noise, $\sqrt{V}$. As for the one-factor model, the partial derivatives of of $log\mathbb{P}$ of the model estimated from bond yields without noise are zero, while the partial derivatives of $logL^{\mathbb{Q}}$ at the true value of the coefficients vary significantly. The partial derivatives of $log\mathbb{Q}$ with respect to $k1$ and $k2$ are large and positive in comparison to partial derivatives with respect the remaining parameters. However, they are about two thirds smaller than for the one-factor model. Adding noise to simulated bond yields does not significantly change the partial derivatives of the log-likelihood of the short rate factor dynamics, but the partial derivatives of $log\mathbb{P}$ with respect to $k1$ and $k2$ present negative mean and median values that are of the order of ten times smaller than the partial derivatives of $log\mathbb{Q}$ with respect to $k1$ and $k2$. The disproportionate sizes of the log-likelihood functions and partial derivatives help to explain why allocating more weight to the physical probability measure in the estimation procedure with low alphas results in a small underestimation of $k1$ and $k2$, and why overestimation is observed for higher values of alpha.

3.4 Conclusion

This paper investigates the finite-sample properties of the time series estimation of the one- and two-factor Cox et al. (1985a,b) models of the term structure. I use Monte Carlo simulations to determine the behaviour of the estimators when introducing measure-scaling weights in the joint log-likelihood function as in Fichtner and Santa-Clara (2012). I estimate the model using a state-space framework, where cross-section errors link observable yields to the state vector, and the implied short rate processes describe the dynamics of the state vector. I reduce the bias found in Ball and Torous (1996) and Phillips and Yu (2005) simply by giving more weight to the likelihood function of pricing errors. The values of the log-likelihood functions and partial derivatives become more balanced. As a consequence, fitting the yield curve and bond option pricing performance of the model greatly improves. When bond yields differ from model-implied values, the CIR model estimation behaves differently for the two-factor case than the on-factor case. For the one-factor model, the bias reduction is still accomplished by giving more weight to the likelihood function of pricing errors. Using two factors, the same tradeoff as in Fichtner and Santa-Clara (2012) is observed. Giving more weight to the likelihood function of pricing errors improves the fitting of the yield curve, while giving more weight to the likelihood function of the state vector dynamics improves option pricing at the expense of the latter. The right choice of the measure-scaling weight, however, depends on the application of the model and how much the simulated dynamics differs from the underlying model-implied dynamics.

3.5 Tables and Figures

Table 3.1: Estimation of one-factor CIR model, $\sqrt{V} = 0$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for one-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to zero basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two-year yield was used to invert for the short rate factor, while the five-year yield is assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The one-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of $100 and a strike price of $82. The initial interest rate factor is 5%.

	k	θ	σ	λ	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Value	0.5000	0.0800	0.0400	-0.1000			3.5411
Alpha = 0.1							
Mean	0.5012	0.0800	0.0393	-0.1010	12013.23	2014.69	3.5230
Median	0.5015	0.0800	0.0382	-0.1009	11928.66	2016.09	3.5226
Std. Dev.	0.0127	0.0014	0.0031	0.0142	2228.26	3.66	0.1167
RMSE	0.0144	0.0014	0.0032	0.0142			0.1172
Alpha = 0.3							
Mean	0.5179	0.0797	0.0364	-0.1160	14152.25	2012.20	3.2688
Median	0.5138	0.0797	0.0363	-0.1115	14207.82	2016.34	3.3367
Std. Dev.	0.0480	0.0017	0.0032	0.0480	303.24	31.98	0.3150
RMSE	0.0512	0.0018	0.0037	0.0505			0.3588
Alpha = 0.5							
Mean	0.5246	0.0795	0.0363	-0.1245	14228.78	1970.64	3.2188
Median	0.5293	0.0798	0.0363	-0.1284	14242.95	2016.26	3.3935
Std. Dev.	0.0851	0.0018	0.0033	0.0832	173.78	170.58	0.4890
RMSE	0.0885	0.0019	0.0037	0.0866			0.5451
Alpha = 0.7							
Mean	0.5512	0.0789	0.0362	-0.1426	14393.62	1795.76	2.8460
Median	0.5387	0.0788	0.0363	-0.1372	14310.11	2006.83	2.9714
Std. Dev.	0.0933	0.0018	0.0033	0.0435	2160.07	349.31	0.6710
RMSE	0.1017	0.0020	0.0038	0.0609			0.9658
Alpha = 0.9							
Mean	0.6052	0.0765	0.0358	-0.1854	21976.34	986.49	2.5434
Median	0.5534	0.0798	0.0363	-0.1416	14291.32	1490.71	3.1743
Std. Dev.	0.4496	0.0053	0.0032	0.3771	16139.32	1584.50	1.1347
RMSE	0.4542	0.0064	0.0044	0.4266			1.3143

Table 3.2: Estimation of one-factor CIR model, $\sqrt{V} = 5$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for one-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to 5 basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two-year yield was used to invert for the short rate factor, while the five-year yield is assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The one-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of \$100 and a strike price of \$82. The initial interest rate factor is 5%.

	k	θ	σ	λ	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Value	0.5000	0.0800	0.0400	-0.1000			3.5411
Alpha = 0.1							
Mean	0.5256	0.0798	0.0369	-0.1117	14225.3073	705.8160	3.3280
Median	0.5174	0.0799	0.0370	-0.1038	14223.7428	706.0691	3.4294
Std. Dev.	0.0361	0.0026	0.0031	0.0336	114.8518	9.7523	0.4694
RMSE	0.0444	0.0027	0.0047	0.0355			0.5193
Alpha = 0.3							
Mean	0.5310	0.0785	0.0368	-0.1604	15264.2243	632.4825	3.0660
Median	0.5253	0.0797	0.0369	-0.1174	14290.4832	697.4732	3.2623
Std. Dev.	0.0918	0.0038	0.0037	0.1421	4622.2472	199.1683	0.7377
RMSE	0.1023	0.0046	0.0050	0.1607			0.8558
Alpha = 0.5							
Mean	0.5582	0.0789	0.0367	-0.1772	17840.9009	425.8544	2.8882
Median	0.5543	0.0791	0.0368	-0.1274	14316.1726	689.1186	3.2013
Std. Dev.	0.2943	0.0050	0.0035	0.2183	8156.8875	665.8846	0.8372
RMSE	0.2986	0.0060	0.0055	0.2504			1.0564
Alpha = 0.7							
Mean	0.5828	0.0774	0.0367	-0.2110	19032.6883	364.1277	2.7578
Median	0.5711	0.0798	0.0364	-0.1464	14260.5071	676.7630	3.0989
Std. Dev.	0.5308	0.0048	0.0031	0.3917	12560.8955	863.4743	1.2478
RMSE	0.5443	0.0055	0.0062	0.3996			1.4721
Alpha = 0.9							
Mean	0.6322	0.0770	0.0367	-0.3483	35965.3695	-3389.3436	2.5792
Median	0.5825	0.0781	0.0366	-0.2068	38212.6483	-1236.6961	2.5831
Std. Dev.	0.9616	0.0053	0.0031	0.9157	19744.6549	5966.4046	1.8160
RMSE	0.9679	0.0109	0.0069	0.9502			2.3684

Table 3.3: Estimation of one-factor CIR model, $\sqrt{V} = 10$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for one-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to 10 basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two-year yield was used to invert for the short rate factor, while the five-year yield is assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The one-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of \$100 and a strike price of \$82. The initial interest rate factor is 5%.

	k	θ	σ	λ	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Value	0.5000	0.0800	0.0400	-0.1000			3.5411
Alpha = 0.1							
Mean	0.5251	0.0809	0.0364	-0.0907	14148.1157	361.9169	3.7322
Median	0.5193	0.0806	0.0364	-0.0809	14150.2331	362.4190	3.6997
Std. Dev.	0.1231	0.0029	0.0001	0.1200	224.6676	9.7101	1.0163
RMSE	0.1279	0.0030	0.0036	0.1203			1.0274
Alpha = 0.3							
Mean	0.5585	0.0790	0.0367	-0.1849	14912.5033	316.1797	3.2477
Median	0.4823	0.0800	0.0368	-0.0986	14207.5491	351.6857	3.6547
Std. Dev.	0.3006	0.0034	0.0006	0.3066	3678.9070	118.4259	1.3455
RMSE	0.3060	0.0035	0.0034	0.3179			1.3907
Alpha = 0.5							
Mean	0.5691	0.0788	0.0366	-0.1590	17775.6649	307.3738	3.2434
Median	0.4609	0.0796	0.0367	-0.1047	15256.3697	343.1609	3.5798
Std. Dev.	0.3508	0.0028	0.0006	0.2991	3070.1516	135.0169	1.7472
RMSE	0.3509	0.0030	0.0034	0.3046			1.7784
Alpha = 0.7							
Mean	0.5961	0.0762	0.0354	-0.3354	20781.4514	-148.9465	2.9015
Median	0.4615	0.0798	0.0361	-0.0953	15282.5828	334.8487	3.7702
Std. Dev.	0.6433	0.0058	0.0013	0.7088	13398.4932	1213.0425	2.7690
RMSE	0.6500	0.0069	0.0047	0.7464			2.9743
Alpha = 0.9							
Mean	0.6921	0.0775	0.0365	-0.2952	21275.9755	-353.6345	2.8697
Median	0.5358	0.0791	0.0366	-0.1556	17321.1602	291.6210	3.0629
Std. Dev.	1.1328	0.0034	0.0011	1.1011	4004.1600	1715.8954	2.7435
RMSE	1.1359	0.0042	0.0036	1.1168			3.8660

Table 3.4: One-factor CIR model Pricing Errors

This table pricing and forecasting root mean-squared errors for the one-factor CIR model with different likelihood weights. The 2-year yield was used to invert for the short rate factors, while the five-year yield is assumed to be measured with error.

Panel A. RMSE (%), $\sqrt{V} = 0$ b.p.								
	1-year	3-year	5-year	7-year	10-year	12-year	15-year	20-year
Alpha = 0.1	0.0018	0.0006	0.0003	0.0011	0.0020	0.0031	0.0039	0.0048
Alpha = 0.3	0.0035	0.0013	0.0015	0.0028	0.0046	0.0070	0.0088	0.0109
Alpha = 0.5	0.0089	0.0040	0.0058	0.0084	0.0127	0.0191	0.0242	0.0302
Alpha = 0.7	0.0189	0.0093	0.0155	0.0201	0.0271	0.0388	0.0500	0.0670
Alpha = 0.9	0.0735	0.0496	0.1259	0.2209	0.3508	0.6193	1.0140	2.2503
Panel B. RMSE (%), $\sqrt{V} = 5$ b.p.								
Alpha = 0.1	0.0782	0.0654	0.0589	0.0558	0.0541	0.0537	0.0531	0.0533
Alpha = 0.3	0.0884	0.0754	0.0690	0.0659	0.0744	0.0841	0.0941	0.1047
Alpha = 0.5	0.0954	0.0799	0.0748	0.0873	0.1079	0.1105	0.1271	0.2085
Alpha = 0.7	0.1076	0.0888	0.1117	0.1244	0.1451	0.1760	0.2016	0.2833
Alpha = 0.9	0.1190	0.1003	0.1257	0.2292	0.3079	0.4145	0.7925	1.4628
Panel C. RMSE (%), $\sqrt{V} = 10$ b.p.								
Alpha = 0.1	0.1585	0.1305	0.1173	0.1116	0.1093	0.1098	0.1112	0.1137
Alpha = 0.3	0.1630	0.1332	0.1236	0.1212	0.1238	0.1337	0.1452	0.1654
Alpha = 0.5	0.1641	0.1344	0.1286	0.1312	0.1400	0.1631	0.1909	0.2462
Alpha = 0.7	0.1658	0.1360	0.1320	0.1380	0.1560	0.1961	0.2431	0.3286
Alpha = 0.9	0.1949	0.1752	0.1286	0.1868	0.2572	0.3943	0.5877	1.1479

Table 3.5: Partial derivatives of the one-factor CIR model

This table reports mean, median, and standard deviation (Std. Dev.) of partial derivatives of the one-factor CIR model likelihood functions at the true value of the coefficients, and at different values for the standard different deviation of the yield noise, $\sqrt{V}$. The partial derivatives are computed from 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two-year yield was used to invert for the short rate factor, while the five-year yield is assumed to be measured with error.

	k	θ	σ	λ
$logL^{\mathbb{P}}$, $\sqrt{V} = 0$ b.p.				
Mean	0.0000	0.0000	0.0000	0.0000
Median	0.0000	0.0000	0.0000	0.0000
Std. Dev.	0.0000	0.0000	0.0000	0.0000
$logL^{\mathbb{Q}}$, $\sqrt{V} = 0$ b.p.				
Mean	39.5698	-1.0616	-2.0858	-1.0059
Median	38.4652	-0.7723	-1.3603	-0.9177
Std. Dev.	12.0477	1.9286	2.9239	3.6010
$logL^{\mathbb{P}}$, $\sqrt{V} = 5$ b.p.				
Mean	-0.7119	0.5676	0.1759	-0.2171
Median	0.6237	-0.1546	0.0842	-0.1091
Std. Dev.	3.6894	2.0457	2.9279	2.7880
$logL^{\mathbb{Q}}$, $\sqrt{V} = 5$ b.p.				
Mean	38.6726	-2.5637	-1.7618	-0.9916
Median	38.8958	-2.8171	-1.8272	-0.8681
Std. Dev.	14.6774	3.6709	2.6487	4.1510
$logL^{\mathbb{P}}$, $\sqrt{V} = 10$ b.p.				
Mean	-0.6362	0.7530	0.2740	-0.5197
Median	0.6835	0.1704	0.0566	0.3133
Std. Dev.	7.8733	2.5987	3.4064	3.5971
$logL^{\mathbb{Q}}$, $\sqrt{V} = 10$ b.p.				
Mean	39.9919	-2.2059	-1.4368	-1.1150
Median	38.8342	-2.3194	-1.0059	-0.8224
Std. Dev.	15.5736	5.3238	3.9123	6.7429

Table 3.6: Estimation of two-factor CIR model, $\sqrt{V} = 0$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for two-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to zero basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two- and ten-year yields were used to invert for the short rate factor, while the one- and seven-year yields are assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The two-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of \$100 and a strike price of \$82. The initial values for the interest rate factors are 0.03 and 0.02 for r_1 and r_2, respectively.

	$k1$	$\theta1$	$\sigma1$	$\lambda1$	$k2$	$\theta2$	$\sigma2$	$\lambda2$	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Value	0.5000	0.0700	0.0400	-0.1000	0.0500	0.0200	0.0400	-0.0200			2.7690
Alpha = 0.1											
Mean	0.5146	0.0693	0.0410	-0.1156	0.0523	0.0204	0.0404	-0.0254	18762.6343	14256.8244	2.7062
Median	0.5116	0.0693	0.0409	-0.1124	0.0516	0.0204	0.0407	-0.0248	18329.2226	14906.9304	2.7442
Std. Dev.	0.0318	0.0010	0.0024	0.0305	0.0102	0.0009	0.0020	0.0071	2620.8347	2455.5069	0.2140
RMSE	0.0350	0.0012	0.0026	0.0342	0.0105	0.0010	0.0020	0.0089			0.2328
Alpha = 0.3											
Mean	0.5396	0.0702	0.0395	-0.1336	0.0721	0.0195	0.0391	-0.0294	23225.5754	2463.9395	2.5801
Median	0.5235	0.0702	0.0389	-0.1206	0.0716	0.0196	0.0389	-0.0266	22269.2355	2165.4284	2.6496
Std. Dev.	0.0805	0.0019	0.0027	0.0852	0.0268	0.0013	0.0008	0.0203	5826.8541	1156.6058	0.5730
RMSE	0.0928	0.0019	0.0030	0.0915	0.0347	0.0014	0.0050	0.0224			0.6273
Alpha = 0.5											
Mean	0.5580	0.0707	0.0396	-0.1801	0.0645	0.0196	0.0387	-0.0284	29035.6208	1934.5918	2.4695
Median	0.5516	0.0707	0.0386	-0.1619	0.0659	0.0199	0.0361	-0.0235	27155.3707	1869.9366	2.4285
Std. Dev.	0.1039	0.0018	0.0024	0.0676	0.0261	0.0015	0.0015	0.0208	7806.1929	682.2654	0.7999
RMSE	0.1553	0.0020	0.0034	0.0705	0.0298	0.0016	0.0055	0.0224			1.0622
Alpha = 0.7											
Mean	0.5855	0.0708	0.0367	-0.1885	0.0619	0.0196	0.0367	-0.0279	32351.0785	1566.4773	2.3390
Median	0.5746	0.0708	0.0367	-0.1779	0.0631	0.0199	0.0341	-0.0239	32343.8148	1689.5310	2.2973
Std. Dev.	0.1678	0.0018	0.0015	0.1146	0.0286	0.0017	0.0016	0.0185	9732.4479	748.3343	1.0824
RMSE	0.2077	0.0020	0.0034	0.1272	0.0310	0.0017	0.0055	0.0201			1.1843
Alpha = 0.9											
Mean	0.6105	0.0710	0.0368	-0.2131	0.0612	0.0195	0.0347	-0.0291	37532.5413	807.7839	2.1522
Median	0.6006	0.0711	0.0367	-0.2055	0.0619	0.0200	0.0341	-0.0240	35359.3716	1470.8196	2.0759
Std. Dev.	0.2043	0.0019	0.0016	0.1548	0.0297	0.0019	0.0017	0.0247	11194.8709	1940.0684	1.5273
RMSE	0.2508	0.0021	0.0033	0.1861	0.0317	0.0019	0.0055	0.0264			1.7941

Table 3.7: Estimation of two-factor CIR model, $\sqrt{V} = 5$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for two-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to 5 basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two- and ten-year yields were used to invert for the short rate factor, while the one- and seven-year yields are assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The two-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of $100 and a strike price of $82. The initial values for the interest rate factors are 0.03 and 0.02 for r_1 and r_2, respectively.

	$k1$	$\theta 1$	$\sigma 1$	$\lambda 1$	$k2$	$\theta 2$	$\sigma 2$	$\lambda 2$	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Values	0.5000	0.0700	0.0400	-0.1000	0.0500	0.0200	0.0400	-0.0200			2.7690
Alpha = 0.1											
Mean	0.4778	0.0692	0.0366	-0.1013	0.0442	0.0190	0.0346	-0.0237	19400.9937	1185.2915	2.9605
Median	0.4879	0.0696	0.0366	-0.1007	0.0452	0.0187	0.0346	-0.0231	12813.9786	1184.3713	2.8932
Std. Dev.	0.0868	0.0023	0.0004	0.0712	0.0179	0.0017	0.0003	0.0171	15194.5830	17.1212	0.7706
RMSE	0.0896	0.0024	0.0034	0.0711	0.0188	0.0020	0.0054	0.0175			0.8869
Alpha = 0.3											
Mean	0.4979	0.0700	0.0387	-0.1081	0.0501	0.0187	0.0347	-0.0252	21031.7934	1132.4610	2.7891
Median	0.5001	0.0701	0.0387	-0.1054	0.0522	0.0186	0.0348	-0.0225	19959.8068	1151.1258	2.7885
Std. Dev.	0.0960	0.0018	0.0012	0.0712	0.0283	0.0014	0.0006	0.0236	14918.1489	71.5041	0.5137
RMSE	0.0959	0.0018	0.0023	0.0731	0.0283	0.0019	0.0053	0.0246			0.5271
Alpha = 0.5											
Mean	0.5385	0.0702	0.0389	-0.1117	0.0570	0.0191	0.0351	-0.0267	27684.4547	995.3861	2.5542
Median	0.5266	0.0702	0.0389	-0.1120	0.0591	0.0192	0.0348	-0.0225	23590.3478	1060.2296	2.5818
Std. Dev.	0.1103	0.0019	0.0013	0.0783	0.0289	0.0015	0.0012	0.0216	12919.0531	230.8388	0.9793
RMSE	0.1204	0.0019	0.0021	0.0811	0.0297	0.0018	0.0050	0.0226			1.0626
Alpha = 0.7											
Mean	0.5707	0.0706	0.0390	-0.1101	0.0587	0.0193	0.0351	-0.0371	32184.0411	562.0891	2.1960
Median	0.5722	0.0706	0.0390	-0.1055	0.0614	0.0195	0.0348	-0.0252	30487.5564	705.4466	2.1820
Std. Dev.	0.1998	0.0020	0.0014	0.0813	0.0312	0.0017	0.0011	0.0424	12202.3450	812.7069	1.0103
RMSE	0.2665	0.0021	0.0020	0.0817	0.0324	0.0019	0.0051	0.0457			1.3058
Alpha = 0.9											
Mean	0.6439	0.0708	0.0393	-0.1104	0.0595	0.0194	0.0352	-0.0324	39594.0501	-261.0499	2.0655
Median	0.6298	0.0708	0.0392	-0.1069	0.0611	0.0197	0.0349	-0.0252	36296.9463	447.3663	2.0780
Std. Dev.	0.2492	0.0016	0.0016	0.0623	0.0274	0.0017	0.0012	0.0345	8923.8880	2472.3087	0.9252
RMSE	0.2987	0.0018	0.0018	0.0630	0.0290	0.0018	0.0049	0.0366			2.2615

Table 3.8: Estimation of two-factor CIR model, $\sqrt{V} = 10$ b.p.

This table reports mean, median, standard deviation (Std. Dev.) and root mean squared errors (RMSE) for two-factor CIR model coefficient estimates, likelihood values and option price, with different likelihood weights and standard deviation of the yield noise, $\sqrt{V}$, set to 10 basis points. All models estimate 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two- and ten-year yields were used to invert for the short rate factor, while the one- and seven-year yields are assumed to be measured with error. Alpha is the likelihood weight of $logL^{\mathbb{Q}}$ in $logL$. The two-factor CIR model was estimated using quasi-maximum likelihood procedure. The one-year European call option on a three-year discount bond has a face value of $100 and a strike price of $82.

	$k1$	$\theta 1$	$\sigma 1$	$\lambda 1$	$k2$	$\theta 2$	$\sigma 2$	$\lambda 2$	$logL^{\mathbb{Q}}$	$logL^{\mathbb{P}}$	Option
True Values	0.5000	0.0700	0.0400	-0.1000	0.0500	0.0200	0.0400	-0.0200			2.7690
Alpha = 0.1											
Mean	0.4735	0.0696	0.0367	-0.0726	0.0325	0.0187	0.0346	-0.0109	21219.8837	390.8646	3.7812
Median	0.4773	0.0689	0.0366	-0.0642	0.0278	0.0179	0.0345	-0.0058	22613.5430	424.9039	3.9506
Std. Dev.	0.1142	0.0042	0.0005	0.1206	0.0281	0.0021	0.0003	0.0302	22304.8257	113.5306	1.6734
RMSE	0.1834	0.0043	0.0034	0.1235	0.0345	0.0025	0.0051	0.0315			2.0407
Alpha = 0.3											
Mean	0.4749	0.0698	0.0368	-0.0848	0.0419	0.0191	0.0346	-0.0281	45376.4033	103.7727	3.2714
Median	0.4836	0.0697	0.0367	-0.0895	0.0423	0.0189	0.0344	-0.0216	33555.7406	190.8387	3.1777
Std. Dev.	0.1249	0.0031	0.0005	0.1066	0.0291	0.0017	0.0005	0.0356	17100.0112	329.0490	1.3667
RMSE	0.1386	0.0031	0.0033	0.1075	0.0303	0.0022	0.0055	0.0365			1.4513
Alpha = 0.5											
Mean	0.5027	0.0700	0.0368	-0.1086	0.0428	0.0194	0.0348	-0.0302	38620.3014	-247.3304	2.7510
Median	0.4998	0.0700	0.0368	-0.1098	0.0436	0.0194	0.0346	-0.0228	32351.8338	-76.3759	2.6595
Std. Dev.	0.1752	0.0024	0.0005	0.0935	0.0302	0.0016	0.0009	0.0350	13826.5049	607.3805	1.1492
RMSE	0.2051	0.0024	0.0032	0.0938	0.0310	0.0019	0.0053	0.0364			1.1403
Alpha = 0.7											
Mean	0.5760	0.0700	0.0370	-0.1165	0.0529	0.0194	0.0350	-0.0372	38335.8493	-791.4567	2.5814
Median	0.5557	0.0702	0.0369	-0.1129	0.0567	0.0197	0.0348	-0.0245	32275.2942	-523.2387	2.4703
Std. Dev.	0.2437	0.0020	0.0005	0.0794	0.0290	0.0016	0.0010	0.0602	14155.5853	1128.9143	1.1099
RMSE	0.3003	0.0020	0.0031	0.0810	0.0291	0.0018	0.0051	0.0626			1.2115
Alpha = 0.9											
Mean	0.6688	0.0703	0.0371	-0.1060	0.0591	0.0199	0.0351	-0.0326	43911.9008	-1890.6700	2.2369
Median	0.6508	0.0703	0.0371	-0.1034	0.0617	0.0202	0.0349	-0.0258	41821.2355	-1562.2954	2.1776
Std. Dev.	0.3745	0.0015	0.0005	0.0693	0.0252	0.0017	0.0009	0.0303	7762.6127	1762.5365	0.9665
RMSE	0.4528	0.0015	0.0029	0.0694	0.0268	0.0017	0.0050	0.0328			1.6022

Table 3.9: Two-factor CIR model Pricing Errors

This table pricing and forecasting root mean-squared errors for the two-factor CIR model with different likelihood weights. The 2 and 10 year yields were used to invert for the short rate factors, while the one- and seven-year yields are assumed to be measured with error.

Panel A. RMSE (%), $\sqrt{V} = 0$ b.p.								
	1-year	3-year	5-year	7-year	9-year	12-year	15-year	20-year
Alpha = 0.1	0.0085	0.0046	0.0078	0.0067	0.0028	0.0072	0.0209	0.0487
Alpha = 0.3	0.0236	0.0118	0.0191	0.0160	0.0066	0.0165	0.0470	0.1068
Alpha = 0.5	0.0362	0.0171	0.0259	0.0209	0.0085	0.0206	0.0585	0.1322
Alpha = 0.7	0.0506	0.0245	0.0372	0.0291	0.0114	0.0265	0.0730	0.1608
Alpha = 0.9	0.0703	0.0338	0.0504	0.0384	0.0146	0.0329	0.0890	0.1922
Panel B. RMSE (%), $\sqrt{V} = 5$ b.p.								
Alpha = 0.1	0.0836	0.0646	0.0627	0.0654	0.0690	0.0743	0.0811	0.0980
Alpha = 0.3	0.0857	0.0649	0.0644	0.0668	0.0693	0.0748	0.0861	0.1190
Alpha = 0.5	0.0910	0.0662	0.0686	0.0698	0.0699	0.0765	0.0965	0.1506
Alpha = 0.7	0.1029	0.0698	0.0769	0.0756	0.0710	0.0800	0.1146	0.1999
Alpha = 0.9	0.1222	0.0754	0.0854	0.0797	0.0716	0.0816	0.1196	0.2109
Panel C. RMSE (%), $\sqrt{V} = 10$ b.p.								
Alpha = 0.1	0.1646	0.1301	0.1250	0.1297	0.1375	0.1499	0.1657	0.2077
Alpha = 0.3	0.1675	0.1307	0.1277	0.1317	0.1376	0.1504	0.1701	0.2197
Alpha = 0.5	0.1726	0.1316	0.1310	0.1345	0.1385	0.1520	0.1801	0.2527
Alpha = 0.7	0.1764	0.1318	0.1335	0.1364	0.1389	0.1519	0.1814	0.2578
Alpha = 0.9	0.1828	0.1330	0.1375	0.1400	0.1401	0.1555	0.2011	0.3740

Table 3.10: Partial derivatives of the two-factor CIR model

This table reports mean, median, and standard deviation (Std. Dev.) of partial derivatives of the two-factor CIR model likelihood functions at the true value of the coefficients, and at different values for the standard different deviation of the yield noise, $\sqrt{V}$. The partial derivatives are computed from 1000 Monte Carlo simulations consisting of 500 weekly observations of the instantaneous interest rate. The two-year yield was used to invert for the short rate factor, while the five-year yield is assumed to be measured with error.

	$k1$	$\theta 1$	$\sigma 1$	$\lambda 1$	$k2$	$\theta 2$	$\sigma 2$	$\lambda 2$
$logL^{\mathbb{P}}$, $\sqrt{V}=0$ b.p.								
Mean	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Median	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Std. Dev.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$logL^{\mathbb{Q}}$, $\sqrt{V}=0$ b.p.								
Mean	13.2139	0.7023	1.8380	-0.5616	8.5547	-1.3675	1.4530	-0.5458
Median	12.3299	0.3696	1.3708	-0.4683	8.8330	-1.0287	1.3378	-0.4763
Std. Dev.	7.5170	3.1217	5.0551	2.3206	4.5761	1.4940	1.9608	1.3230
$logL^{\mathbb{P}}$, $\sqrt{V}=5$ b.p.								
Mean	-1.5918	-0.5481	-0.3607	-0.1365	-1.0567	0.3513	0.0165	0.1973
Median	-1.3364	-0.1017	-0.1143	0.0907	-1.3312	0.0339	0.0945	0.0617
Std. Dev.	5.2187	2.4376	1.2722	4.8164	2.2700	1.9154	1.0745	2.6654
$logL^{\mathbb{Q}}$, $\sqrt{V}=5$ b.p.								
Mean	13.5821	2.0278	1.4905	-0.5454	9.5918	-1.3330	-1.9616	-0.5224
Median	12.7627	1.7884	1.0486	-0.5364	8.8178	-1.5682	-1.9319	-0.5504
Std. Dev.	8.5434	2.9640	6.9471	2.1516	5.9751	1.3984	1.6143	1.8246
$logL^{\mathbb{P}}$, $\sqrt{V}=10$ b.p.								
Mean	-2.3006	-0.4962	-0.6210	-0.1901	-1.0128	-0.0726	0.0833	0.3938
Median	-2.0019	-0.2033	-0.0749	-0.0574	-1.3208	0.0677	-0.0060	0.2116
Std. Dev.	6.0265	2.8753	1.6922	5.0200	3.9973	2.8309	1.3626	3.6214
$logL^{\mathbb{Q}}$, $\sqrt{V}=10$ b.p.								
Mean	14.3421	1.6400	1.6042	-0.5869	9.5051	-1.9565	-1.4308	-0.7746
Median	13.8086	1.4101	1.0350	-0.5079	8.8334	-1.2962	-1.1914	-0.5466
Std. Dev.	9.5615	5.1997	4.0864	2.2611	7.2634	1.0761	1.6493	2.4208

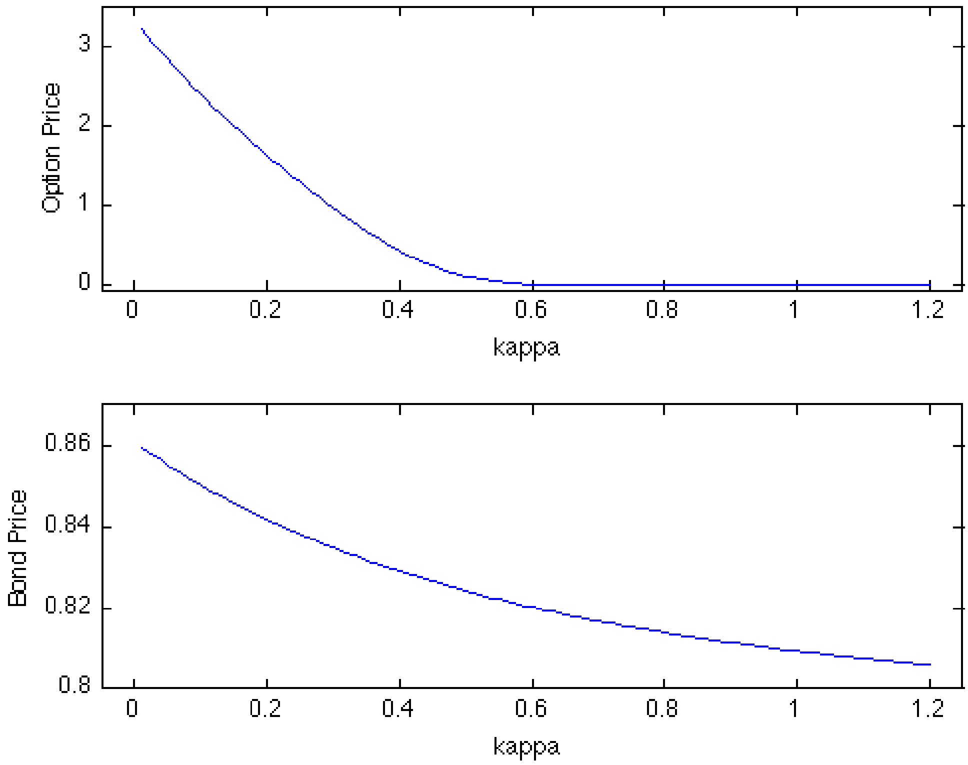


Figure 3.1: Relationship between k and option and bond prices

This shows the price of a three-year zero coupon bond and the price of one-year in-the-money European call option on a three-year zero coupon bond as a function of the speed of adjustment parameter, k, under a CIR model. The assumed parameters are $\theta = 0.08$, $\sigma = 0.02$ and $\lambda = 0$. The face value of the coupon bond is \$1 and the initial interest rate is 5%. The option on the discount bond has a face value of \$100 and a strike price of \$87.

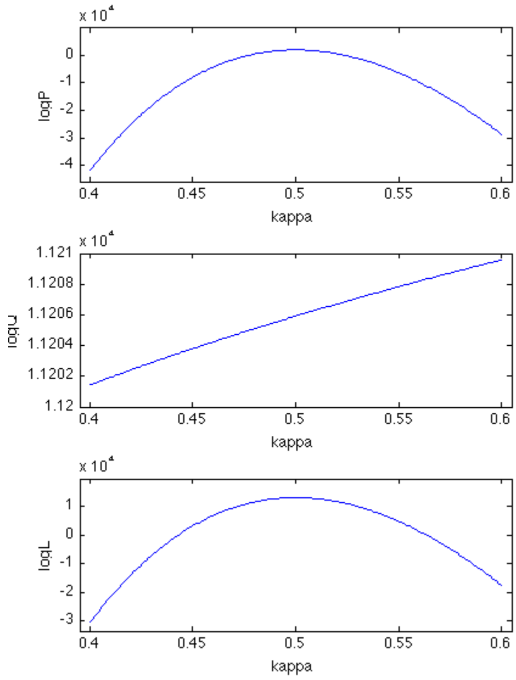


Figure 3.2: Relationship between k and log-likelihood values

This shows the log-likelihood values corresponding to the log-likelihood of pricing errors, $logL^{\mathbb{P}}$, the log-likelihood of pricing errors, $logL^{\mathbb{Q}}$, and the joint log-likelihood, $logL$ as a function of the speed of adjustment parameter, k, under a one-factor CIR model. The yield curve was simulated for 500 weekly observations without measurement errors. The assumed parameters are $\theta = 0.07$, $\sigma = 0.03$ and $\lambda = -0.1$. Alpha is set at 0.5. The likelihood values of $logL^{\mathbb{P}}$, $logL^{\mathbb{Q}}$ and $logL$ using $k = 0.5$ are 1.9708e+03, 1.1206e+ 04 and 6.5884e+03, respectively. At $k = 0.5$, $\partial logL^{\mathbb{Q}}/\partial k = 40.1395$, and $\partial logL^{\mathbb{P}}/\partial k = 0$.

Bibliography

Ait-Sahalia, Yacine, "Transition Densities for Interest Rate and Other Nonlinear Diffusions," *The Journal of Finance*, 1999, *54* (4), pp. 1361–1395. 25, 30, 31

_ , "Closed-Form Likelihood Expansions for Multivariate Diffusions," *The Annals of Statistics*, 2008, *36* (2), pp. 906–937. 25, 30, 31

_ **and Robert L. Kimmel**, "Estimating affine multifactor term structure models using closed-form likelihood expansions," *Journal of Financial Economics*, 2010, *98* (1), 113 – 144. 24, 30, 39, 47, 48, 49, 53

Ashley, Richard A., "Beyond Optimal Forecasting," Working Papers e06-10, Virginia Polytechnic Institute and State University, Department of Economics 2006. 2

Backus, David, Silverio Foresi, Abon Mozumdar, and Liuren Wu, "Predictable changes in yields and forward rates," *Journal of Financial Economics*, 2001, *59* (3), 281 – 311. 23

Ball, Clifford A. and Walter N. Torous, "Unit roots and the estimation of interest rate dynamics," *Journal of Empirical Finance*, 1996, *3* (2), 215 – 238. ii, 24, 34, 48, 51, 59

Bansal, Ravi and Hao Zhou, "Term Structure of Interest Rates with Regime Shifts," *The Journal of Finance*, 2002, *57* (5), pp. 1997–2043. 23

Black, Fischer, "The pricing of commodity contracts," *Journal of Financial Economics*, 1976, *3* (12), 167 – 179. 28

Brace, Alan, Dariusz Gatarek, and Marek Musiela, "The Market Model of Interest Rate Dynamics," *Mathematical Finance*, 1997, *7* (2), 127–155. 24

Campbell, John Y. and Robert J. Shiller, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *The Review of Economic Studies*, 1991, *58* (3), pp. 495–514. 1, 2

Chapman, David A. and Neil D. Pearson, "Is the Short Rate Drift Actually Nonlinear?," *The Journal of Finance*, 2000, *55* (1), pp. 355–388. 34

_ **and** _ , "Is the Short Rate Drift Actually Nonlinear?," *The Journal of Finance*, 2000, *55* (1), pp. 355–388. 48

Chen, Ren-Raw and Louis Scott, "Pricing Interest Rate Options in a Two-Factor Cox–Ingersoll–Ross Model of the Term Structure," *The Review of Financial Studies*, 1992, *5* (4), pp. 613–636. 24, 29, 30, 52

_ **and** _ , "Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates," *Journal of Fixed Income*, 1993, *3* (3), pp. 14–31. 24, 25, 30, 31, 35, 48, 53, 54

_ **and** _ , "Interest rate options in multifactor Cox-Ingersoll-Ross models of the term structure," *Journal of Derivatives*, 1995, *3* (2), pp. 53–72. 29

Christensen, Jens H.E., Francis X. Diebold, and Glenn D. Rudebusch, "The affine arbitrage-free class of Nelson-Siegel term structure models," *Journal of Econometrics*, 2011, *164* (1), 4 – 20. Annals Issue on Forecasting. 24

Cochrane, John H., *Asset Pricing*, Princeton University Press, January 2005. 9

_ **and John Campbell**, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," Scholarly Articles 3119444, Harvard University Department of Economics 1999. 8

_ **and Monika Piazzesi**, "Bond Risk Premia," *The American Economic Review*, 2005, *95* (1), pp. 138–160. 1, 2, 3, 9, 10, 11, 13

Constantinides, George M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 1990, *98* (3), pp. 519–543. 8

Cox, John C., Jr. Ingersoll Jonathan E., and Stephen A. Ross, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 1985, *53* (2), pp. 363–384. ii, 23, 24, 37, 47, 52, 59

_ , _ , and _ , "A Theory of the Term Structure of Interest Rates," *Econometrica*, 1985, *53* (2), pp. 385–407. ii, 23, 24, 25, 27, 37, 47, 52, 54, 59

Dai, Qiang and Kenneth J. Singleton, "Specification Analysis of Affine Term Structure Models," *The Journal of Finance*, 2000, *55* (5), pp. 1943–1978. 23, 24, 27, 47, 52

de Jong, Frank, "Time-series and cross section information in affine term structure models," *Journal of Business and Economic Statistics*, 2000, *18* (3), pp. 300–314. 31, 48

Diebold, Francis X. and Canlin Li, "Forecasting the term structure of government bond yields," *Journal of Econometrics*, 2006, *130* (2), 337 – 364. 3, 6, 11, 13, 24

Duffee, Gregory R., "Term Premia and Interest Rate Forecasts in Affine Models," *The Journal of Finance*, 2002, *57* (1), pp. 405–443. 24, 31

_ and Richard H. Stanton, "Estimation of Dynamic Term Structure Models," Technical Report 2004. 24, 25, 34, 47, 48, 49, 54, 55

Duffie, Darrell and Kenneth J. Singleton, "An Econometric Model of the Term Structure of Interest-Rate Swap Yields," *The Journal of Finance*, 1997, *52* (4), pp. 1287–1321. 24, 27, 30, 48, 54

_ and _ , "Modeling Term Structures of Defaultable Bonds," *The Review of Financial Studies*, 1999, *12* (4), pp. 687–720. 26

_ and Rui Kan, "A Yield-Factor Model Of Interest Rates," *Mathematical Finance*, 1996, *6* (4), 379–406. 23, 27, 47

Fama, Eugene F. and Robert R. Bliss, "The Information in Long-Maturity Forward Rates," *The American Economic Review*, 1987, *77* (4), pp. 680–692. 1, 2, 3, 6, 9, 11, 13

Feller, William, "Two Singular Diffusion Problems," *The Annals of Mathematics*, 1951, *54* (1), pp. 173–182. 27

Ferreira, Miguel A. and Pedro Santa-Clara, "Forecasting stock market returns: The sum of the parts is more than the whole," *Journal of Financial Economics*, 2011, *100* (3), 514 – 537. 7

Fichtner, Luiz P. and Pedro Santa-Clara, "How You Estimate the Yield Curve Matters!," Technical Report 2012. 48, 49, 50, 51, 53, 54, 56, 58, 59

Gibbons, Michael R. and Krishna Ramaswamy, "A Test of the Cox, Ingersoll, and Ross Model of the Term Structure," *The Review of Financial Studies*, 1993, *6* (3), pp. 619–658. 48

Goyal, Amit and Ivo Welch, "Predicting the Equity Premium with Dividend Ratios," *Management Science*, 2003, *49* (5), pp. 639–654. 2, 7, 8

Greenwood, Robin and Dimitri Vayanos, "Bond Supply and Excess Bond Returns," CEPR Discussion Papers 6694, C.E.P.R. Discussion Papers February 2008. 1, 2, 9, 10, 13

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright, "The U.S. Treasury yield curve: 1961 to the present," Finance and Economics Discussion Series 2006-28, Board of Governors of the Federal Reserve System (U.S.) 2006. 3

Hull, J.C., *Options, Futures, and Other Derivatives*, 7th ed., Pearson Prentice Hall, 2008. 28

Ilmanen, Antti, "Time-Varying Expected Returns in International Bond Markets," *The Journal of Finance*, 1995, *50* (2), pp. 481–506. 1, 2, 8, 10, 13

Jagannathan, Ravi, Andrew Kaplin, and Steve Sun, "An evaluation of multifactor CIR models using LIBOR, swap rates, and cap and swaption prices," *Journal of Econometrics*, 2003, *116* (1), 113 – 146. Frontiers of financial econometrics and financial engineering. 24, 25, 27, 30, 48, 53, 54

Jamshidian, Farshid, "LIBOR and swap market models and measures (*)," *Finance and Stochastics*, 1997, *1* (4), 293–330. 24

Lamoureux, Christopher G. and H. Douglas Witte, "Empirical Analysis of the Yield Curve: The Information in the Data Viewed through the Window of Cox, Ingersoll, and Ross," *The Journal of Finance*, 2002, *57* (3), pp. 1479–1520. 24, 25, 27, 30, 48, 53, 54

Longstaff, Francis A. and Eduardo S. Schwartz, "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance*, 1992, *47* (4), pp. 1259–1282. 24

_ , Pedro Santa-Clara, and Eduardo S. Schwartz, "The Relative Valuation of Caps and Swaptions: Theory and Empirical Evidence," *The Journal of Finance*, 2001, *56* (6), pp. 2067–2109. 24

Longstaff, Francis A, Pedro Santa-Clara, and Eduardo S Schwartz, "Throwing away a billion dollars: the cost of suboptimal exercise strategies in the swaptions market," *Journal of Financial Economics*, 2001, *62* (1), 39 – 66. 24

Miltersen, Kristian R., Klaus Sandmann, and Dieter Sondermann, "Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates," *The Journal of Finance*, 1997, *52* (1), pp. 409–430. 24

Orphanides, Athanasios and Don H. Kim, "Term Structure Estimation with Survey Data on Interest Rate Forecasts," Computing in Economics and Finance 2005 474, Society for Computational Economics November 2005. 24, 34

Pearson, Neil D. and Tong-Sheng Sun, "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model," *The Journal of Finance*, 1994, *49* (4), pp. 1279–1304. 24, 25, 27, 48, 54

Phillips, Peter C. B. and Jun Yu, "Jackknifing Bond Option Prices," *The Review of Financial Studies*, 2005, *18* (2), pp. 707–742. ii, 24, 25, 26, 34, 36, 48, 51, 53, 54, 55, 59

Siegel, Andrew F. and Charles R. Nelson, "Long-Term Behavior of Yield Curves," *The Journal of Financial and Quantitative Analysis*, 1988, *23* (1), pp. 105–110. 6

Vasicek, Oldrich, "An equilibrium characterization of the term structure," *Journal of Financial Economics*, 1977, *5* (2), 177 – 188. 23, 47

Yu, Jun and Peter C. B. Phillips, "A Gaussian approach for continuous time models of the short-term interest rate," *Econometrics Journal*, 2001, *4* (2), pp. 211–225. 34, 48

Erratum to Three Essays on the Term Structure of Interest Rates

Luiz Paulo Fichtner

May 31, 2013

To the members of the jury:

Following the attention raised Raquel M. Gaspar on the confusing notation used in Chapters 2 and 3 of the Doctoral Thesis entitled "Three Essays on the Term Structure of Interest Rates" concerning likelihood function and attribution of probability measures, I would like to ask the members of the jury to consider the following change in notation: where it reads "log-likelihood function of cross-section pricing errors under the physical measure" and "log-likelihood function of the state vector dynamics under the equivalent risk-neutral measure" (page 25, line 2, 3, 12, 25; page 31, line 7, 11; page 32, line 14, page 33, line 29, 30; page 37, line 20, 21; page 47, line 15 to 18; page 49, line 12, 13; page 50, line 27; page 53, line 18, 21; page 54, line 8; page 55, line 7, 8, 16, 22, 23; page 56, line 13, 32; page 57, line 21, 23; page 58, line 21), please consider only "log-likelihood function of cross-section pricing errors" and "log-likelihood function of the state vector dynamics". In the course of writing the Thesis the notation was altered and usage of measures related to the likelihood functions derived from a need to simplify the notation. However, it comes in contradiction with the parameters used to compute these likelihood functions. Therefore the usage of $logL^{\mathbb{Q}}$ and $logL^{\mathbb{P}}$ is also mixed (specially in Equations 2.9 and 3.7, pages 32 and 53, and the tables of results). They should rather be simply understood as $logL^{Dynamics}$ and $logL^{Cross-section}$ instead.

This wrong usage of notation, however, did not influence the estimation of the CIR model in Chapters 2 and 3. The parameters used to compute cap prices, yield forecasts and yield pricing errors are correctly used in the Thesis and computation of results (expect for typos in Equation

2.5 and 2.6, page 28, where it should be kbar instead of k).

In light of this, it also came to my attention that the parameters used in the computation of option prices in Chapter 3 are wrong due to my inaccurate usage of a Matlab program from Phillips and Yu (2005), where the market price of risk is not considered. Therefore the option prices are wrongly computed. I would like to ask the jury to please disconsider this result.

However, like mentioned above, the estimation of the CIR model in this chapter is not influenced, and nor are the computation of partial derivatives and yield pricing errors. This is an issue that I will have to correct in the future version of the Chapter.

I thank Raquel M. Gaspar for voicing this issue, and ask the jury for the understanding about the errors contained in my PhD Thesis.

My sincere apologies,

Luiz Paulo Fichtner