Review of Finance (2005) 9: 97–137

Interpreting Implied Risk-Neutral Densities: The Role of Risk Premia ⋆

PETER HÖRDAHL and DAVID VESTIN
European Central Bank

Abstract. This paper examines differences between risk-neutral and objective probability densities of future interest rates. The identification and quantification of these differences are important when risk-neutral densities (RNDs), such as option-implied RNDs, are used as indicators of actual beliefs of investors. We employ a multi-factor essentially affine modeling framework applied to German time-series and cross-section term structure data in order to identify both the risk-neutral and the objective term structure dynamics. We find important differences between risk-neutral and objective distributions due to risk premia in bond prices. Moreover, the estimated premia vary over time in a quantitatively significant way, which implies that the differences between the objective and the risk-neutral distributions also vary over time. We therefore conclude that one should be cautious in interpreting RNDs in terms of expectations. The method used in this paper provides an alternative approach to identifying objective probabilities of future interest rates.

1. Introduction

The interplay between financial markets and central banks provides rich opportunities for both market participants and monetary policy makers to extract valuable information from financial asset prices. Market participants monitor and forecast the policy decisions of central banks in order to price interest-rate related contracts and other financial assets. Conversely, central banks are interested in evaluating the markets expectations about its future interest-rate policy and about future underlying fundamentals, such as growth prospects and inflation. Prices of financial contracts (e.g., bonds, futures and options) are an obvious source for extracting this kind of information and a large number of techniques have been developed towards achieving this end, including the modelling of implied forward rates and implied

⋆ We are grateful for useful suggestions from Greg Duffee, John Fell, Nikolaos Panigirtzoglou, Neil Pearson, Ranjini Sivakumar, an anonymous referee, participants at the ECB Workshop on Measures and Determinants of Financial Market Uncertainty in Frankfurt, the 2003 European Finance Association meetings in Glasgow, the 2003 European Financial Management Association meetings in Helsinki, and the 2003 Econometric Society European Meeting; seminar participants at the Swedish Riksbank, the Oesterreichische Nationalbank, and Lehman Brothers in London. We would also like to thank Darrell Duffie for very useful discussions on topics of relevance for this paper. The views expressed herein are those of the authors and do not necessarily correspond to those of the European Central Bank or the Eurosystem.

volatilities. Another approach which has been increasingly used to examine measures of uncertainty, is to model the entire implied density of the price of a financial contract for a future date of interest.

This paper focuses on the role of risk premia when examining and interpreting the information from such financial indicators in the context of interest rate-related markets. This issue has, in practice, often been disregarded or, alternatively, an implicit assumption has often been made that the impact of premia is negligible. One example is the relatively common practice to interpret implied forward rates as direct measures of expected future interest rates, without explicitly accounting for risk premia considerations.

In the context of implied densities, the existence of possible premia may in principle also have important implications for the interpretation of the information from such densities. A common approach when estimating implied densities is to assume some specific parametric specification for the density, and then calibrate the parameters to minimize pricing errors with respect to a given cross-section of observed derivatives prices with identical expiration. However, it is well-known that this approach will deliver the so-called risk-neutral density of the underlying price or yield at the time of expiration, and not the actual objective density. More precisely, derivatives pricing relies largely on the absence of arbitrage, rather than on some "objective" valuation theory. The absence of arbitrage, in turn, (essentially) implies the existence of an equivalent martingale measure, often referred to as the "risk neutral measure", which is the relevant probability measure to use when pricing derivatives as the discounted expected payoff. Heuristically, this risk-neutral pricing framework implies that the probability measure used to price assets is adjusted so as to make the expected return on a risky asset equal to the risk free rate. It does not, however, mean that agents are assumed to be risk neutral.

The fact that option-implied densities deliver the risk-neutral probabilities is not a problem if the aim is to price other securities, because this is the relevant density for such purposes. It may, however, be a problem if the aim is to interpret market expectations of uncertainty surrounding future interest rates or asset prices, since the relevant distribution in this case of course is the true, objective density. This paper aims at exploring the differences between these measures, in order to find out whether one can safely ignore risk premia considerations when interpreting RNDs, which largely seems to be current practice. Unfortunately, it turns out to be difficult to analyze the impact of risk premia when the RND is modelled directly. One possibility is to assume some specific functional form for the utility function and then estimate the degree of risk aversion in order to back out the true PDF from the RND, as in Bliss and Panigirtzoglou (2004). However, an alternative approach is available, in that we can follow the pricing literature that models the dynamic properties of the underlying asset. A few studies have taken a similar route in examining risk-neutral vs. objective PDFs for equity indices, but these have focused on quite restrictive single-factor specifications, as e.g., in the study by Aït-Sahalia, Wang and Yared (2001). In contrast, we are mainly interested in examining these issues in connection with interest-bearing instruments, for which a single-factor setup is likely to be even less realistic than in the case of equities. We therefore turn to recently proposed dynamic multi-factor term structure models. We focus on the class of affine term structure models (ATSMs) of Duffie and Kan (1996) and Dai and Singleton (2000), which have received increasing attention in the term structure literature due to their flexibility and analytical tractability. In particular, within this class of models, Duffee (2002) has shown that a setup which allows risk premia to depend on the factors in a flexible way does well in capturing the dynamics of premia, and therefore also improves yield forecasts. Hence, we employ this so-called "essentially affine" specification in our analysis of implied densities. In this context, it is interesting to note that Egorov et al. (2003) provide empirical evidence that the essentially affine specification of Duffee (2002) does well in terms of forecasting U.S. yield densities.

Once a model for the term structure dynamics is specified, the distributional properties of yields can be uniquely pinned down. Moreover, given that we specify the functional form of the risk premium, Girsanov's theorem allows us to quantify the differences between the risk-neutral and the real-world probability measures. In particular, we can explore the differences in the means of the two distributions for arbitrary combinations of maturities and horizons, and examine if and how these differences vary over time. This is, for example, of interest from a classical expectations hypothesis point of view when one is interested in expected future interest rates. In addition, our framework allows us to investigate how various measures of dispersion differ between the risk-neutral and the objective probability measures. This relates to the issue of interpreting the dispersion of implied densities as a measure of the market's perceived degree of uncertainty with respect to the future evolution of the underlying. Apart from the impact on simple variance measures, we can also quantify the differences between the measures in terms of the probabilities for various scenarios, such as "the probability that the 10-year yield will be lower than $x\%$ in y months". This is of interest because implied distributions are often implicitly or explicitly used to get an idea of such probabilities.

We focus our analysis mainly on the results for distributions on 3-month interest rates and 10-year yields for horizons of up to one year, because these will be the most relevant cases in practice. Specifically, the most liquid standardized interest rate options markets in the euro area are the markets for options on three-month EURIBOR futures and for options on 10-year German government bond futures (Bunds). The results from this exercise indicate that risk premia considerations are important, in the sense that there generally are non-negligible differences between risk-neutral and objective densities. More importantly, we find that the differences between these densities change over time as a result of time-variation in risk premia.

The remainder of the paper is organized as follows. Section 2 reviews the theoretical foundation for the two alternative modelling strategies: the direct density modelling approach and the dynamic process modelling approach. Section 3 illustrates the latter approach using the well-known and simple Cox, Ingersoll and Ross (1985) model (CIR), while Section 4 considers more general affine multi-factor models. Section 5 provides some density forecast evaluations, and section 6 concludes.

2. Theory

In the absence of arbitrage opportunities it is possible to show, given technical conditions, that the price of any contingent claim, Π at time t, based on an underlying asset Y at time T, will be given by

$$\Pi\left(t, Y\right) = E_t^Q\left[e^{-\int_t^T r(s)ds}\phi\left(Y_T\right)\right], \tag{1}$$

where $\phi\left(Y_T\right)$ is the payoff of the claim and $r\left(s\right)$ is the stochastic interest rate. This is the fundamental result on which both the option-implied approach and the dynamic process approach relies. Equation (1) states that the price of the claim, say a European call option, will be given by the discounted value of the payoff of the claim. The crucial thing to note is that the expected value should be evaluated under an equivalent martingale measure, often denoted the risk-neutral measure (Q). This is the source of the term risk-neutral density (RND).

This section reviews two ways of modelling this risk-neutral density. First, we examine the case when the distribution is modelled directly. This line of work is represented by the Melick and Thomas (1997) approach, which is widely used by central banks to infer interest rate distributions from market prices. However, as argued above, central bankers are mainly interested in the objective probability measure, which differs from the risk-neutral measure due to risk premia. Therefore, implicit in the interpretation of option-implied densities is an assumption that risk premia do not matter "too much". Alternatively, it is assumed that such premia remain fairly constant over time, so that *changes* in the risk neutral density can be interpreted in terms of changes in the actual, objective density. Of course, the main appeal of the direct approach is that it often is relatively easy to implement in order to obtain some indication of the distribution of the underlying. On the other hand, a weakness with the direct approach to modelling the distribution is that it is impossible to assess the importance of risk premia, since the analysis is conducted under the risk-neutral measure.

To evaluate the quantitative differences, we therefore introduce a second approach, which amounts to modelling the evolution of the underlying state variables that drive the dynamics of the yield curve. Specifically, we rely on the affine modelling approach of Duffie and Kan (1996) and Dai and Singleton (2000), whereby the underlying state variables and the instantaneous short-term interest rate follow affine diffusion processes. An early and classical example of a model within this framework is Cox, Ingersoll and Ross (1985). In the affine framework, it is possible to relatively easily specify the relation between the risk-neutral and the objective probability measures, and to use market data to estimate all parameters governing the dynamics under the different measures. Thus, if both approaches properly capture the risk-neutral density, and the assumptions about the functional form of the risk premium are correct, then we can examine the quantitative errors associated with approximating the objective density with the risk-neutral one. We elaborate some more on the two approaches mentioned above, in the following two sub-sections.

We start by introducing some definitions. Let $P(t, T)$ denote the price, at t, of a zero-coupon bond maturing at T. Correspondingly, $y(t, T)$ is the yield to maturity for such a bond,

$$y(t, T) = -\frac{\ln(P(t, T))}{T - t} \tag{2}$$

2.1. MODELLING THE DISTRIBUTION

A specific case of (1) which we will be interested in is European options. In this case, the price of a call option at t can be expressed as

$$C(t, T; K, P) = E_t^Q\left[e^{-\int_t^T r(s)ds}[P(T) - K]^+\right] \tag{3}$$

where $P(T)$ is the value of the underlying contract (e.g., a bond) at the expiration date T of the option, and K is the strike price. In the case of futures options with full margining, the discount effect vanishes and the option price formula collapses to

$$C(t, T; K, P) = E_t^Q[P(T) - K]^+ .$$

This is relevant, because standardized interest rate options contracts are typically options on futures, as in the case of EURIBOR and Bund futures options.

Now, if we are prepared to assume that the risk-neutral distribution of $P(T)$ is given by $f^Q(x)$, then we can use market data to get hold of the parameters of the distribution. For example, assume that $f^Q(x)$ is a mixture of two log-normal distributions, as suggested by Melick and Thomas (1997). Then, the risk-neutral density is given by

$$f^Q(x) = \theta L(x; \mu_1, \sigma_1) + (1 - \theta) L(x; \mu_2, \sigma_2), \tag{4}$$

where

$$L(x, \mu_i, \sigma_i) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu_i)^2}{2\sigma_i^2}}.$$

Let $\tilde{C}(t, T; K, \mu, \sigma)$ denote the theoretical call option price obtained by integrating the above distribution for a given set of parameters, $\gamma = [\mu_1, \mu_2, \sigma_1, \sigma_1, \theta]'$. Assuming that we, for a given expiration date T, have market data on a set of option prices with different strikes, indexed by i, we can obtain estimates of the parameters by solving

$$\min_{\gamma} \sum_{i} \left(C(t, T; K_i, P) - \tilde{C}(t, T; K_i, \gamma) \right)^2 . \tag{5}$$

Further, if one is prepared to assume that the Q-measure coincides with (or is approximately equal to) the objective measure, the graph of $f^Q(x)$ can be interpreted in terms of the uncertainty perceived by the market about the future price of the underlying, e.g., a bond. Taking the example of a bond price density, with an appropriate density transformation, this can be converted into a statement about the distribution of future yields.[1] Our interest is therefore to evaluate how large the differences between the risk-neutral and the objective densities are. It is conceptually hard to do this in the approach outlined above. Nevertheless, as the next section will show, there is a natural relationship between the two measures when using an alternative modelling strategy.

2.2. MODELLING THE TERM STRUCTURE

Most of the bond pricing literature models the dynamic evolution of a number of state variables or factors, X, that drive the prices of interest-related contracts. This is typically done by specifying the functional form of the deterministic and the diffusion terms, together with initial conditions for the state variables

$$\begin{aligned} dX(t) &= \mu(t, X(t))\, dt + \sigma(t, X(t))\, dW(t), \\ X(t) &= X_t, \end{aligned} \tag{6}$$

where $W(t)$ is a vector-valued Brownian motion, specified under the objective probability-measure P. For notational convenience we will from this point on use X_t, W_t to denote $X(t)$, $W(t)$, etc.

For reasons of tractability, it is useful to assume that zero-coupon bond prices are exponential-affine functions of the state variables, that is

$$P(t, T) = e^{A(t,T) + B(t,T) \cdot X_t}, \tag{7}$$

where $A(t, T)$ and the factor loadings $B(t, T)$ are functions of the time to maturity $T - t$, with initial conditions $A(0) = B(0) = 0$. It turns out that the requirements needed for this to apply is that $\mu(t, X_t)$ and $\sigma(t, X_t)\sigma'(t, X_t)$ are affine functions of the state variables under the Q-measure. In this case, we also know that the

[1] More specifically, the Jacobian of the transformation is given by differentiating (2).

instantaneous interest rate r will be affine in the state variables, so that $r_t = \delta_0 + \delta_1 X_t$. Once the objective dynamics are specified, Girsanov's theorem provides the link between the two measures. Specifically, if W is a (potentially vector-valued) Wiener process under one measure, then it follows from Girsanov's theorem that W^Q is a Wiener process under the Q measure,

$$dW_t^Q = dW_t - g_t dt, \tag{8}$$

where g_t is the kernel in the transformation. The so-called "market price of risk", Λ_t is defined to be the negative of the Girsanov kernel, g_t (see Duffie (2001) for further details).

Restricting our attention to the affine class of models, by applying Ito's formula on (7) and imposing the restrictions of no-arbitrage, it is possible to show that the required local rate of return on any bond is equal to the instantaneous interest rate r plus a factor which is proportional to the local volatility of the bond return $\sigma_t (P(t, T))$,

$$\mu_t (P(t, T)) = r_t + \Lambda_t \sigma_t (P(t, T)).$$

From this relation, it is easy to see intuitively why Λ_t is commonly referred to as the market price of risk. Rearranging, we see that Λ is the required rate of return in excess of the risk-free short rate, divided by the volatility,

$$\Lambda_t = \frac{\mu_t (P) - r_t}{\sigma_t (P)}. \tag{9}$$

Moreover, it is clear that the market price of risk is required to be identical irrespective of the maturity of the bond in order to preclude arbitrage opportunities.

Once the functional form of μ, σ and Λ have been specified, and the parameters are estimated using market (panel) data, we can in principle solve (6) and calculate the distributions of future interest rates. Two things can be noted. First, if the model is affine under Q, it is possible to easily calculate prices of bonds and contingent claims. Second, if the model is affine under P, estimation of the model using time-series data is facilitated. Current practice in the literature is to assume that the functional form of the market price of risk is such that both these properties hold. Dai and Singleton (2000) provide an overview of this affine framework, whereas Duffee (2002) suggests an "essentially affine model", meaning that while the factor dynamics are affine under both P and Q, the variance of the state price deflator is not affine. This latter approach has proven to be important with respect to the forecasting abilities of the model. We will return to this issue in some more detail in Section 4.

It is interesting to compare the framework discussed above to the situation in the Black and Scholes (1973) model for equity options. In that model, there is a unique transformation between P and Q because markets are complete. In our case, this is no longer true. From a mathematical perspective, markets are incomplete in the sense that the exogenously specified risk sources are not traded assets and bonds can therefore not be replicated by taking positions in those assets (in clear contrast to the Black and Scholes model). However, once the functional form of the risk premium has been chosen, the correspondence between the martingale measure and the objective measure is uniquely pinned down. At least in principle, all that is required to completely identify the measure is time series observations on the same number of bonds as there are risk sources in the model. With long enough time series, it should be possible to identify the correct functional form of the risk premium. Thus, in a sense, markets are (informationally) complete from the perspective of the econometrician.

3. The Cox, Ingersoll and Ross Model

In order to take a closer look at the issues discussed in previous sections, and at the same time keep the discussion at a relatively intuitive level, we first provide an illustration based on a simple affine one-factor model, namely the well-known Cox, Ingersoll and Ross (CIR, 1985) model. In this model, the term structure is completely determined by the dynamic behavior of one state variable which, in turn, can be expressed in terms of the dynamics of the short-term interest rate,

$$dr_t = \kappa\,(\theta - r_t)\,dt + \sigma\sqrt{r_t}dW_t, \tag{10}$$

where dW_t is a standard Brownian motion increment. This equation describes the interest rate dynamics under the objective probability measure P. It is clear from this specification that the short rate r will revert towards a constant level θ at an adjustment rate determined by κ, and that the variance of the short rate is proportional to the interest rate level.

Given the assumptions underlying their model, CIR obtain closed-form solutions for bond prices (P) of any given maturity. These prices depend on the current value of the state variable (the instantaneous interest rate), the parameters of the model, the time to maturity of the bond, and the market's required compensation for bearing risk. In the CIR case, the local expected return can be shown to be equal to $r + \lambda r P_r/P$, where P_r denotes the partial derivative of the bond price with respect to the short rate r, and λr is the covariance of changes in the interest rate with changes in optimally invested wealth (see CIR (1985), p. 393). The compensation for risk is therefore $\lambda r P_r/P$, which will be positive if λ is negative, since $P_r < 0$. Based on the discussion in the previous section, we know that the market price of risk must satisfy the condition (9) for bonds of any maturity. By substituting for the local expected rate of return and the volatility in the CIR model, we can identify the specific form for the market price of risk in the CIR model as

$$\Lambda_{CIR} = \frac{\lambda\sqrt{r}}{\sigma}. \tag{11}$$

Equipped with the market price of risk, we can now obtain the dynamics of the instantaneous interest rate r under the risk-neutral, or equivalent martingale, measure Q using Girsanov's theorem, whereby

$$dr_t = \tilde{\kappa}\left(\tilde{\theta} - r_t\right)dt + \sigma\sqrt{r_t}d\tilde{W}_t, \tag{12}$$

where $d\tilde{W}_t$ is a Brownian motion increment under Q, and

$$\tilde{\kappa} = \kappa + \lambda, \tag{13}$$

$$\tilde{\theta} = \frac{\kappa\theta}{(\kappa + \lambda)}, \tag{14}$$

are the speed and level of mean reversion under the risk-neutral measure Q. Hence, given a fixed set of parameters for the interest rate process under P, a larger negative value for the risk parameter λ will imply a higher level $\tilde{\theta}$ towards which the short rate will revert under Q, as well as a slower speed of adjustment than under the actual probability measure. In contrast, we see that the instantaneous volatility of r remains unchanged after the change of probability measure. However, this does not mean that the volatility over a discrete time interval will be unchanged, since it will be dependent on the parameters in the drift specification as well. This is easily seen from the analytical expression for the variance of $r(s)$ conditional on $r(t)$, $s > t$, in the CIR model[2],

$$\begin{aligned}\mathrm{Var}\left[r(s) \mid r(t)\right] = r(t)\left(\frac{\sigma^2}{\kappa}\right)\left(\exp\left[-\kappa(s-t)\right] - \exp\left[-2\kappa(s-t)\right]\right) \\ + \theta\left(\frac{\sigma^2}{2\kappa}\right)\left(1 - \exp\left[-\kappa(s-t)\right]\right)^2.\end{aligned} \tag{15}$$

The impact on the conditional variance of a change in the risk parameter can be divided into two components. First, a larger (negative) λ leads to a slower speed of mean-reversion under Q, (13), which in turn results in a higher variance because the short rate r is pulled back towards the long run mean at a slower pace relative to P. Second, a larger (negative) λ increases the risk neutral mean-reversion level, (14), thereby raising the variance (since it is increasing in the level of the short rate).

Given the differences between the short-term interest rate processes under the real-world probability measure P and the risk-neutral measure Q, it is of interest to examine what they imply for the distribution of bond yields under P and Q. The conditional density of the short term interest rate is available in closed form, as is the corresponding conditional distribution function, which is non-central chi-square (see CIR (1985, pp. 391–392)). Moreover, since bond yields, y, for

[2] See Cox et al. (1985, p. 392).

Table I. GMM estimates of CIR parameters under the objective probability measure P, and fitted value of λ

The parameter values in the table are obtained by estimating the discretized CIR model: $\Delta r_t = \kappa(\theta - r_t)\Delta t + \sigma\sqrt{r_t}\Delta W_t$ on German weekly data between January 1996 and March 2002. Standard errors are asymptotic, based on the optimal weighting matrix of Hansen (1982).

Parameter	Estimate	St. err.
κ	0.523	0.257
θ	0.031	0.005
σ	0.027	0.003
λ	−0.295	–

maturities longer than instantaneous simply are affine functions of the short rate, $y(\tau) = a(\tau) + b(\tau)r(t)$, their conditional densities are easily obtained through a transformation of the analytical short-rate density. The functions $a(\tau)$ and $b(\tau)$ depend on time to maturity (τ) as well as on the parameters of the interest rate process and the market price of risk (see CIR (1985) for the explicit formulae).

In order to illustrate, we use the parameter values in Table I to calculate densities under P and under Q. The values in Table I were obtained by estimating the model with the methods used by Aït-Sahalia (1996) on German weekly term structure data between January 1996 and March 2002.

Figure 1 displays the distribution of the short rate $r(s)$ conditional on $r(t)$ for the case when the horizon $(s - t)$ is one year, and when the initial short rate is equal to 3.4%.[3] From the figure, a couple of observations can immediately be made. First, the risk-neutral density lies to the right of the actual density. This follows from the previously mentioned fact that for a negative value of the risk parameter λ, the short-term interest rate will revert towards a higher level under Q than under the actual probability measure P. Second, the risk-neutral densities are more dispersed compared to the actual ones, as discussed above.

Since yields are affine in the instantaneous interest rate, the densities for longer-maturity yields will retain the functional form of the short rate density in Figure 1. The specific shape of a density for a given maturity τ will, apart from the value of the state variable, depend on the function $b(\tau)$, which in turn depends on the parameters governing the risk-neutral dynamics, as well as on time to maturity itself. Moreover, the differences between the P and Q densities for any given horizon will depend on the differences between the physical and risk-neutral dynamics of r, and on the forecast horizon. In the simple setting of the CIR model, the price of risk parameter λ will, apart from a Jensen's inequality term, determine how big these differences turn out to be.

[3] The initial short rate 3.4% used in this example is the level of the shortest interest rate at the end of the data sample.

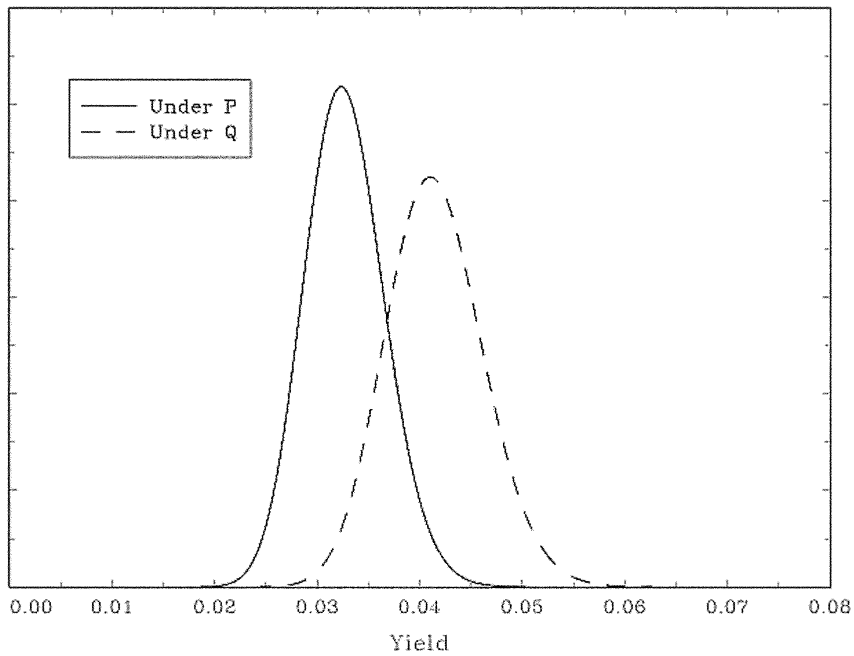

Figure 1. Short-term interest rate distribution under P and Q, 1 year ahead, for initial state variable $r(t) = 3.4\%$.

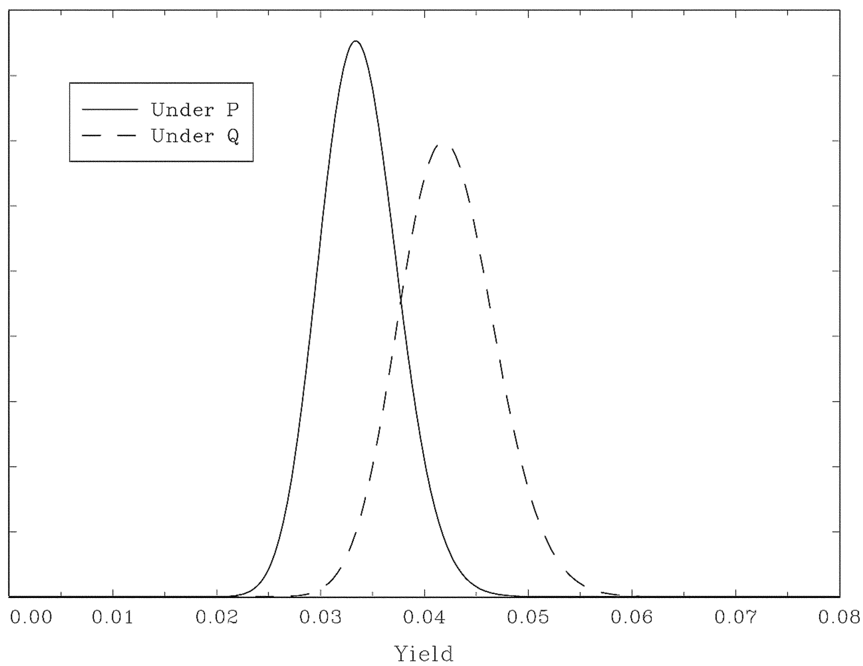

Figure 2. Three-month interest rate distribution under P and Q, 1 year ahead, for initial state variable $r(t) = 3.4\%$.

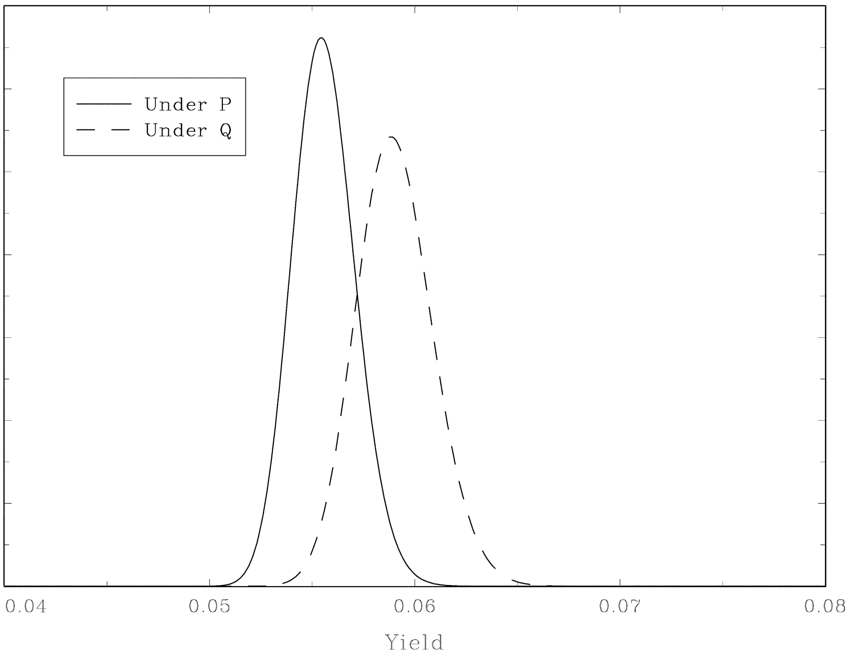

Figure 3. Ten-year yield distribution under P and Q, 1 year ahead, for initial state variable $r(t) = 3.4\%$.

Figure 2 displays the one-year ahead distribution of the three-month interest rate conditional on an initial short rate of 3.4%. It is not surprising that the three-month densities resemble the instantaneous densities in Figure 1, given the relative similarity of the maturities. Greater differences are visible for the corresponding one-year ahead densities of ten-year bond yields in Figure 3. First, the densities are centered around higher levels than the densities in Figures 1 and 2, reflecting the upward-sloping tendency of the yield curve. Second, the long-term yield densities are substantially less dispersed than the shorter-rate densities. This result is due to the fact that the term structure of yield volatilities tends to be downward-sloping; i.e., long-term yields tend to be less volatile than short-term rates (on an annualized basis). In terms of holding returns, however, densities based on long-term bonds would be more dispersed than return densities based on short-term bonds.

4. Three-factor Essentially Affine Models

Moving away from the one-factor CIR case, we now turn to multi-factor models which provide more flexibility in capturing various features of term structure data over time. To this end, we focus on the affine class of dynamic term structure models which has attracted increasing attention by researchers as well as practitioners, in particular following Duffie and Kan's (1996) presentation of a generalized affine modelling framework. Empirical research has shown that three factors seem sufficiently flexible to capture important aspects of the dynamics of term structure data (e.g., De Jong (2000), Dai and Singleton (2002), Duffee (2002)). While more factors in theory should provide additional flexibility, practical considerations tend to rule out estimation of high-dimensional dynamic term structure models. We therefore focus on three-factor affine models in this section, in the hope that these will provide a reasonable trade-off between flexibility and analytical tractability.

Given our choice to use a three-factor affine specification to model the term structure, a large number of different variants are available within this class, depending on how we choose to parameterize the model. Recently, Dai and Singleton (2000) has proposed a classification of multi-factor ATSMs into subfamilies of admissible models depending on the choice of variance specification.[4] Specifically, each N-factor ATSM can be categorized into $N + 1$ subfamilies, according to the number of factors that drive the conditional factor variances.

Consider a general N-factor ATSM where the instantaneous short-term interest rate is an affine function of the factors X_t,

$$r_t = \delta_0 + \delta_x' X_t, \tag{16}$$

and where X_t follows an affine diffusion,

$$\begin{aligned} dX_t &= \mathcal{K}\,(\Theta - X_t)\,dt + \Sigma\sqrt{S_t}dB_t \\ &\equiv \mu\,(X)\,dt + \sigma\,(X)\,dB_t, \end{aligned} \tag{17}$$

where dB_t is an N-dimensional vector of independent standard Brownian motions under the objective probability measure P, while $\mathcal{K}$ and Σ are $N \times N$ parameter matrices, and Θ is an N-vector of parameters. The $N \times N$ matrix S_t is diagonal, with diagonal elements given by

$$[S_t]_{ii} = \alpha_i + \beta_i' X_t, \tag{18}$$

where α_i is a scalar and β_i is $N \times 1$. For future reference, we let α denote the N-vector consisting of the individual α_i:s, and

$$\mathcal{B} \equiv (\beta_1 \;\; \beta_2 \;\; \beta_3)$$

denote the $N \times N$ matrix of coefficients on X_t. By imposing restrictions on the matrix $\mathcal{B}$, we can obtain different versions of the N-factor model with respect to the degree of dependence of the conditional variances on the factors. The categorization by Dai and Singleton (2000) identifies $N + 1$ general subfamilies in this

[4] Dai and Singleton (2000) refer to an ATSM as admissible if it ensures positive factor variances, or, more precisely, if if it ensures that $[S_t]_{ii}$ (see below) is strictly positive for each i (see their Appendix B for further details). Their classification scheme imposes the minimal known sufficient conditions for admissibility and and the minimal normalizations for econometric identification.

regard, where the variance $\sigma(X)$ is driven by m factors and m can take values from 0 to N. Specifically, letting $m \equiv \text{rank}(\mathcal{B})$, Dai and Singleton use $\mathbb{A}_m(N)$ to denote the set of admissible N-factor ATSMs with m factors determining the factor variance matrix.

For any given specification of an N-factor ATSM, a number of invariant transformations can be made, in which the state and parameter vectors undergo various transformations and rescalings, while resulting in an unchanged instantaneous short rate and unchanged bond prices (see Appendix A in Dai and Singleton (2000)). In order to facilitate the task of checking whether some given model specification is admissible, they propose a specific invariant transformation, the so-called canonical representation of admissible ATSMs, which imposes minimum constraints to ensure admissibility and econometric identification. Basically, the canonical representation partitions the factors into a vector consisting of the m factors (if any) which drive the conditional variances, and the $N - m$ other factors, so that $X' = \left(X^{B\prime}_{m\times 1}, X^{D\prime}_{(N-m)\times 1}\right)$. The parameter matrices and vectors are then partitioned accordingly, and the following normalizations are imposed:

$$\mathcal{K} = \begin{bmatrix} \mathcal{K}^{BB}_{m\times m} & 0_{m\times(N-m)} \\ \mathcal{K}^{DB}_{(N-m)\times m} & \mathcal{K}^{DD}_{(N-m)\times(N-m)} \end{bmatrix}, \tag{19}$$

if $m > 0$, and with $\mathcal{K}$ triangular otherwise,

$$\Theta = \begin{pmatrix} \Theta^{B}_{m\times 1} \\ 0_{(N-m)\times 1} \end{pmatrix}, \tag{20}$$

$$\Sigma = I, \tag{21}$$

$$\alpha = \begin{pmatrix} 0_{m\times 1} \\ 1_{(N-m)\times 1} \end{pmatrix}, \tag{22}$$

and

$$\mathcal{B} = \begin{bmatrix} I_{m\times m} & \mathcal{B}^{BD}_{m\times(N-m)} \\ 0_{(N-m)\times m} & 0_{(N-m)\times(N-m)} \end{bmatrix}. \tag{23}$$

Furthermore, a number of parameter restrictions are imposed in order to assure admissibility and econometric identification (see Dai and Singleton (2000) for details).

4.1. A GAUSSIAN THREE-FACTOR ATSM: $\mathbb{A}_0(3)$

Starting off with the simplest three-factor canonical ATSM, consider the case where none of the factors affect the volatility of X_t, i.e., an $\mathbb{A}_0(3)$ model. In this case, the canonical representation sets α equal to a 3×1 vector of ones, while $\mathcal{B}$

is set to zero, so that $\sigma(X) = I_3$. Hence, we are left with a homoskedastic model, where the factors follow a Gaussian diffusion. Clearly, interest rates and bond yields do not have constant variances, but this model can nevertheless serve as a simple initial point of departure. Moreover, as shown by Duffee (2002), it turns out that there is a trade-off between the flexibility of the variance specification and the possibility to specify a flexible market price of risk. Furthermore, the flexibility of the market price of risk specification has proved crucial for improving interest rate forecasts and for capturing important features in the data (Duffee (2002), Dai and Singleton (2002)). Consequently, the "simple" Gaussian model may not necessarily fare as badly as one could fear a priori.

The key to the arguments above regarding the trade-off between the flexibility of the variance specification and the market price of risk, is that one moves away from the "standard" assumption for the market price of risk specification, which has that

$$\Lambda(t) = \sqrt{S_t}\lambda, \tag{24}$$

where λ is a vector of constants. In this setup, compensation for risk will be proportional to the variance of the risk factors. The advantage of this simple specification is that it preserves the affine structure when changing measure from the objective to the risk-neutral probability measure.[5] Duffee (2002), on the other hand, proposes a specification for the market price of risk which breaks the link between compensation for risk and the variance.[6] The resulting class of models, denoted the "essentially affine" class, allows not only the variance of the factors to determine the risk compensation, but also lets the factors themselves influence the compensation for risk. It turns out that the essentially affine specification also preserves the affine features of the model under both the objective and the risk-neutral measure.[7]

In more detail, the essentially affine class defines the market price of risk vector as

$$\Lambda_t = \sqrt{S_t}\lambda + \sqrt{S_t^-}\psi X_t, \tag{25}$$

where $\sqrt{S_t^-}$ is defined as the following diagonal matrix

$$\sqrt{S_{t(ii)}^-} = \begin{cases} \left(\alpha_i + \beta_i' X_t\right)^{-1/2} & \text{if } \inf\left(\alpha_i + \beta_i' X_t\right) > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{26}$$

[5] ATSMs with this type of market price of risk are therefore sometimes referred to as "completely affine".

[6] Duarte (2004) proposes an alternative specification ("semi-affine") which also provides added flexibility in the market price of risk. However, this specification does not in general allow greater flexibility than that of Duffee (2002). Moreover, the semi-affine class is not affine under the objective probability measure, thereby making it computationally more burdensome to handle in practice.

[7] The label "essentially affine" refers to the fact that while the factor processes remain affine under a change of measure, the variance of the state price deflator, $\Lambda\Lambda'$, is not affine. This is, however, inconsequential since this variance does not affect bond prices.

while ψ is a 3×3 parameter matrix. Given the canonical representation as specified in (19)-(23), it is easy to see why there is a trade-off between the flexibility of $\sigma(X)$ and that of Λ. Clearly, in the $\mathcal{A}_3(3)$ case, where by (22) $\alpha = 0$ and by (23) $\mathcal{B}$ is non-zero, $(\alpha_i + \beta_i' X_t)$ can reach zero for each of the three factors i, if X_t were to reach zero. The matrix $\sqrt{S_{t(ii)}^{-}}$ is therefore defined to be zero according to (26), and Λ collapses to the completely affine specification (24). In contrast, in the Gaussian $\mathcal{A}_0(3)$ model, $\alpha_i = 1$ and $\beta_i = 0$ for all i, which means that $\sqrt{S_t^{-}} = \sqrt{S_t} = I_3$, and ψX_t therefore fully impacts on the market price of risk. In between these two extreme cases, $m = 1$ or 2 will imply models with more flexible variance specifications and less flexible Λ than in the $\mathcal{A}_0(3)$ case, but still more flexible risk prices than in the $\mathcal{A}_3(3)$ specification.

As mentioned above, the degree of flexibility in the risk price has proven essential with respect to the ability of ATSMs to predict future interest rates. However, while the $\mathcal{A}_0(3)$ model allows the most flexible specification of Λ, it remains to be seen how successful it is in pricing derivative contracts, since the accurate pricing of such contracts typically require that the model is able to successfully capture the dynamics of the term structure of volatilities.

In order to evaluate the performance of the $\mathcal{A}_0(3)$ model, and anticipating future applications to German Bund futures options, we estimate the parameters of the model using German term structure data from January 1983 to March 2002. Specifically, we obtain a monthly time-series of parameter estimates for Svensson's (1994) extension of the Nelson and Siegel (1987) model from the BIS. These parameters, in turn, have been obtained by fitting the model to the prevailing German yield curve at the end of each month, i.e., to available money market and government bond data. These parameters allows us to obtain zero-coupon bond prices and yields for any maturity every month during the sample period, i.e., a time-series of German term structures. Since the introduction of the euro in January 1999, we can also view these as proxies for the euro area term structures.

The estimation of ATSMs is based on a Kalman filter technique, where the data used consists of zero-coupon yields for the maturities 3, 6, 9, and 12 months, as well as 2, 3, 5, 7, and 10 years. Based on the parameter estimates obtained from the BIS, the estimated Nelson-Siegel-Svensson yield curves at times display a somewhat erratic behavior at the very short end of the curve, in particular in the early part of the sample. We therefore substitute the model-based 3-month rates with a series of actual observed 3-month DEM (EUR after December 1998) interbank interest rates. Figure 4 displays the data used in the estimations.

4.2. ESTIMATION USING THE KALMAN FILTER

The estimation of multi-factor ATSMs is complicated by the fact that there is typically no closed-form solution available for the density of discretely sampled yields or bond prices which could be used in maximum likelihood (ML) estima-

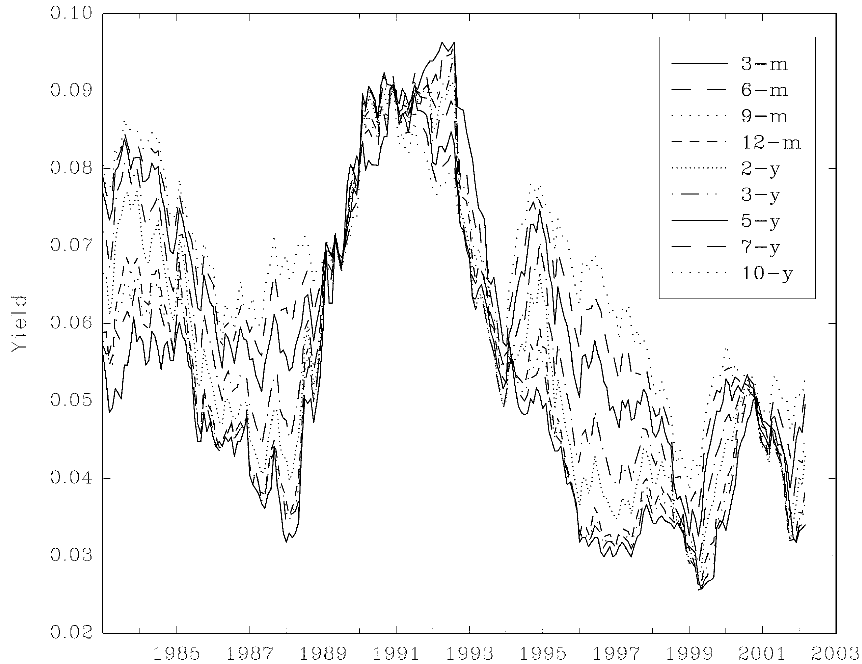

Figure 4. Yield data used in the estimations of multi-factor ATSMs.

tion. However, for Gaussian models, the conditional density *is* known, and we can therefore proceed with ML estimation in the $\mathcal{A}_0\,(3)$ case. Typically in empirical applications, the number of maturities used in the estimation is set equal to the number of factors (e.g., Dai and Singleton (2000)), or, alternatively, it is assumed that yields for N maturities are observed without any error whereas some error structure is imposed on any additional maturities included in the estimation (e.g., Duffee (2002)). In this way, the unobservable factors can be inverted for using the assumed perfectly observable yields. One problem with this approach is that is that it is not entirely clear why the yields for some more or less arbitrarily chosen maturities should be observed perfectly, whereas other yields are observed with some measurement error.

Instead, we follow among others Lund (1997), de Jong (2000), and Duffee and Stanton (2001) in assuming that all yields are observed imperfectly, and applying the Kalman filter technique to estimate the underlying unobservable factors. In principle, these "measurement errors" can be seen as reflecting e.g., bid-ask spreads, non-synchronous data, or other market-specific influences. In this state-space setup, the evolution of the factors over discrete time intervals determines the transition equation, while the measurement equation is taken to be the relation between yields of different maturities and the factors. Moreover, the prediction errors from the Kalman filter and their associated covariances can be used to obtain the exact log-likelihood function in the Gaussian case, or to construct a quasi log-likelihood function in more general cases.

Assume that the data set consists of a time series of length T of M zero-coupon bond yields with constant maturities, $Y_t = (y_t(\tau_1), \ldots, y_t(\tau_M))$, where

$$y_{it} = -\frac{\ln P_{it}}{\tau_i},$$

and τ_i is the time to maturity of bond i. Let the observed yields be equally spaced over time, at intervals of length h. Furthermore, in line with the arguments above, assume that the observed yields consist of the theoretical yields obtained from some chosen model and a measurement error. This provides us with the measurement equation,

$$Y_t = d(\phi) + Z(\phi) X_t + \varepsilon_t, \tag{27}$$

where the measurement error ε_t is assumed to be normally distributed,

$$\varepsilon_t \sim N(0, H(\phi)), \tag{28}$$

and ϕ denotes the vector of parameters in the model. In practice, we assume that $H(\phi)$ is diagonal, and include the diagonal elements of this matrix in the set of parameters to be estimated. Comparing the measurement equation (27) with the theoretical expression for the yield on a zero-coupon bond with maturity τ in the affine framework,

$$y_t(\tau) = -\frac{\ln P_t(\tau)}{\tau} = -\frac{1}{\tau}\left(A(\tau) + B(\tau)' X_t\right),$$

we see that

$$d(\phi) = -\frac{A(\tau)}{\tau} \tag{29}$$

and

$$Z(\phi) = -\frac{B(\tau)'}{\tau}. \tag{30}$$

As mentioned before, the functions $A(\tau)$ and $B(\tau)$ are solutions to a system of ODEs. Specifically, in the $\mathcal{A}_0(3)$ case, we have

$$\frac{dA(\tau)}{d\tau} = -\delta_0 - B'(\tau)\lambda + \frac{1}{2}B'(\tau)B(\tau), \tag{31}$$

$$\frac{dB(\tau)}{d\tau} = -\delta_y - \psi' B(\tau) - \mathcal{K}' B(\tau). \tag{32}$$

Here, we can note that since $A(\tau)$ and $B(\tau)$ are functions of, among other things, the λ and ψ parameter matrices in the market price of risk (25), we are able to estimate these parameters and thereby recover the risk-neutral dynamics.

Next, the transition equation describes the evolution of the state vector from one observation time to the next,

$$X_t = \Phi(\phi) X_{t-h} + u_t, \tag{33}$$

where

$$\Phi(\phi) = \exp[-\mathcal{K}h], \tag{34}$$

and $\exp[-\mathcal{K}(h)]$ refers to the matrix exponential function, defined as

$$\exp[-\mathcal{K}h] = \sum_{k=0}^{\infty} \frac{1}{k!} (-\mathcal{K}h)^k .$$

The transition equation (33) does not contain any constant in the $\mathcal{A}_0(3)$ case since the normalization (20) in the canonical representation sets $\Theta = 0$. Because the model is Gaussian, the innovation in the transition equation is normally distributed,

$$u_t \sim N(0, V_t(\phi)), \tag{35}$$

where the covariance matrix is given by (see e.g., Lund (1997) or Duffee (2002))

$$V_t(\phi) = \int_{t-h}^{t} \exp[-\mathcal{K}(t-s)] \exp\left[-\mathcal{K}'(t-s)\right] ds. \tag{36}$$

Given the specification of the measurement and the transition equations, the Kalman filter algorithm can be implemented to provide a sequence of predictions and updates of the state vector and its variance. Furthermore, the likelihood function is obtained as a result of the Kalman filter recursions, thereby enabling parameter estimation. Appendix A contains some further technical details on the implementation of the Kalman filter in this setting.[8]

Applying the estimation procedure outlined above, we obtain parameter estimates for the $\mathcal{A}_0(3)$ model, as reported in Table II. Since we are able to estimate the price of risk parameters, we have effectively identified both the objective real-world dynamics and the risk-neutral dynamics of the underlying factors, as postulated by the $\mathcal{A}_0(3)$ specification. This enables us to fully characterize the features of the short-term interest rate or zero-coupon bond yields of any maturity, including the conditional distribution of these variables under both probability measures. Hence, assuming that the specification of the model is correct, we now have a

[8] See also e.g., Lund (1997) or de Jong (2000) for further details on implementing the Kalman filter to estimate ATSMs.

Table II. QML parameter estimates for the essentially affine $\mathcal{A}_0\,(3)$ model

The standard errors are based on the asymptotic variance-covariance matrix of White (1982). The estimates of the measurement error variances in $H\,(\phi)$ are not reported.

Parameter	Estimate	Standard error
δ_0	0.052	0.002
$\delta_1 \times 10^2$	0.037	0.178
$\delta_2 \times 10^2$	0.554	0.085
$\delta_3 \times 10^2$	0.926	0.079
κ_{11}	0.196	0.140
κ_{21}	−0.415	0.116
κ_{22}	1.363	0.204
κ_{31}	−0.314	0.334
κ_{32}	1.344	0.494
κ_{33}	0.151	0.080
λ_1	−0.405	0.219
λ_2	−0.052	0.032
λ_3	−1.469	0.267
ψ_{11}	−0.191	0.102
ψ_{12}	0.765	0.365
ψ_{13}	0.009	0.019
ψ_{21}	0.185	0.169
ψ_{22}	−0.151	0.186
ψ_{23}	−0.059	0.021
ψ_{31}	0.431	0.162
ψ_{32}	0.374	0.451
ψ_{33}	0.026	0.021

way of examining the discrepancies between risk-neutral interest rate distributions, which is what an option-implied approach would deliver, and objective interest rate distributions, which is what policymakers are interested in.

The $\mathcal{A}_0\,(3)$ specification allows us to recover yield distributions very easily, since we know that the conditional yield distribution will be normally distributed. Specifically, the mean of the distribution of a τ-maturity zero-coupon yield at time $t+h$ conditional on information at time t will be

$$\begin{aligned} E\left[y_{t+h}\,(\tau)\mid X_t\right] &= -\frac{A\,(\tau)}{\tau} - \frac{B\,(\tau)'}{\tau} E\left[X_{t+h}\mid X_t\right] \\ &= -\frac{A\,(\tau)}{\tau} - \frac{B\,(\tau)'}{\tau}\exp\left[-\mathcal{K}h\right] X_t, \end{aligned} \tag{37}$$

where the second equality follows from (33) and (34). Moreover, the corresponding conditional variance will be given by

$$\begin{aligned}
&\operatorname{Var}\left[y_{t+h}(\tau) \mid X_t\right] \\
&= \frac{B(\tau)'}{\tau} \operatorname{Var}\left[X_{t+h} \mid X_t\right] \frac{B(\tau)}{\tau} \\
&= \frac{B(\tau)'}{\tau} \left(\int_t^{t+h} \exp\left[-\mathcal{K}(t-s)\right] \exp\left[-\mathcal{K}'(t-s)\right] ds \right) \frac{B(\tau)}{\tau}. \qquad (38)
\end{aligned}$$

Hence, the conditional objective (P) distribution of $y_{t+h}(\tau)$ will be Gaussian, with mean given by (37) and variance (38). The corresponding risk-neutral (Q) distribution can be found by using the risk-neutral dynamics when evaluating the conditional expectation and variance of X :

$$E^Q\left[X_{t+h} \mid X_t\right] = \left(I - \exp\left[-\mathcal{K}^Q h\right]\right) \Theta^Q + \exp\left[-\mathcal{K}^Q h\right] X_t, \qquad (39)$$

$$\operatorname{Var}^Q\left[X_{t+h} \mid X_t\right] = \int_t^{t+h} \exp\left[-\mathcal{K}^Q(t-s)\right] \exp\left[-\mathcal{K}^{Q\prime}(t-s)\right] ds, \qquad (40)$$

where $\mathcal{K}^Q = (\mathcal{K} + \psi)$, and $\Theta^Q = -(\mathcal{K} + \psi)^{-1} \lambda$ follow from Girsanov's theorem.

To illustrate the distinction between the objective and the risk-neutral distributions, consider the one-year ahead conditional distributions of the 3-month and the 10-year zero-coupon yields. Figures 5 and 6 display these distributions for each of the two probability measures, as implied by the parameter estimates and the filtered factors at the end of our sample (end-March 2002).

While the differences in the case of the 3-month rate seem relatively small for the particular factor values in this example, there are non-negligible differences between the risk-neutral and the objective densities in the case of the 10-year yield. In the example shown, the mean of the objective distribution is about 25 basis points higher than in the risk-neutral one, while the yield standard deviation is about 0.7% in the P-distribution, and around 0.85% in the Q-distribution.[9] More importantly, the implied probabilities for different yield outcomes can differ substantially. For example, in the example above, the risk-neutral probability that the 10-year yield would be below 5.0% one year ahead is approximately 18.6%, whereas the actual P-probability for the same outcome is only around 7.8%.

In addition, the conditional means of the yield distributions will change over time, as the factors evolve. The relation between the objective and the risk-neutral means will also change, and therefore the relation between the probabilities for

[9] To put the mean difference in perspective, over the entire sample between 1983 and 2002 the estimated average difference in the absolute value of the difference between the means is 34 basis points.

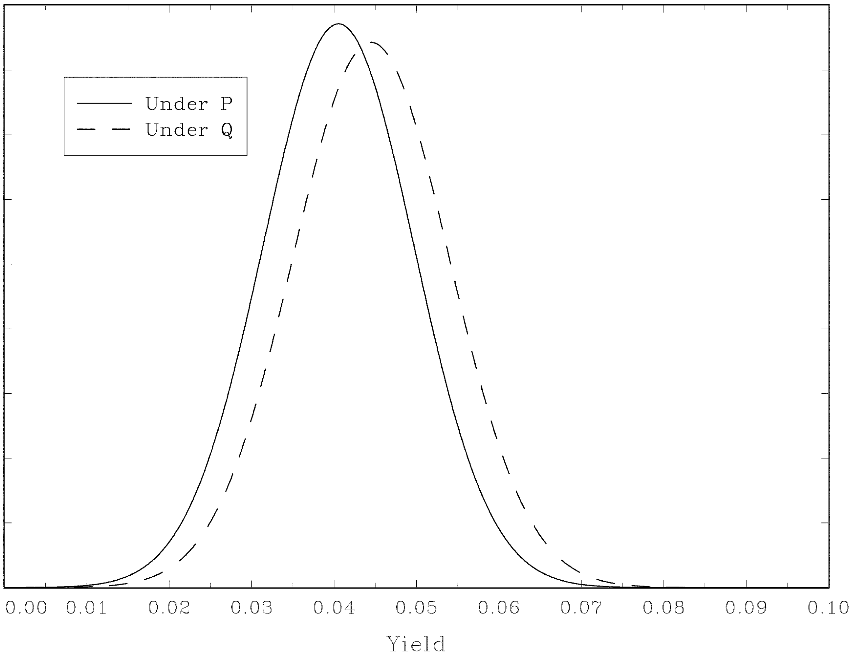

Figure 5. One-year ahead conditional distribution of the 3-month interest rate, as implied by the $\mathcal{A}_0$ (3)-model, end-March 2002.

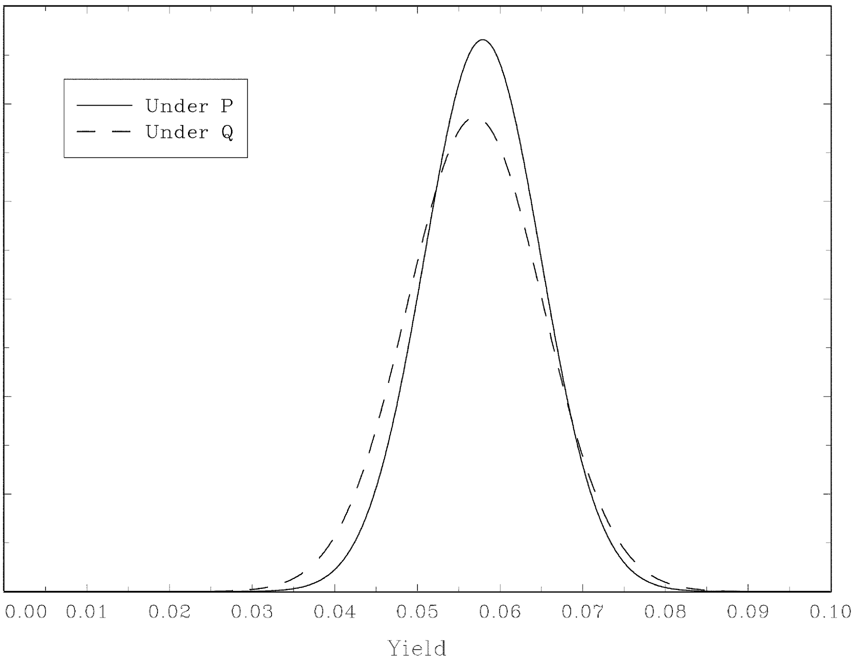

Figure 6. One-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the $\mathcal{A}_0$ (3)-model, end-March 2002.

different outcomes will change. In fact, depending on the evolution of the factors, the difference between the means of the distributions may change sign during some periods. However, it is easy to see from (36) that the variances of the P and Q distributions will remain constant over time, despite variation in the factors. In order to allow greater flexibility in the conditional variance specification, we turn to the $\mathbb{A}_1(3)$ model in the next section.

4.3. THE $\mathbb{A}_1(3)$ ESSENTIALLY AFFINE MODEL

In the $\mathbb{A}_1(3)$ model, only one of the three factors affect the volatility of X_t. In accordance with the canonical representation, this factor is assumed to be the first one, and the parameter vectors and matrices are adjusted according to (19)–(23). Notably, the first element of the vector of long-run means, Θ, is no longer zero, whereas $\alpha_1 = 0$, and the first row of $\mathcal{B}$ is different from zero. We impose suitable parameter restrictions to ensure admissibility and identification (see Dai and Singleton (2000)).

The implications of the $\mathbb{A}_1(3)$ specification on the conditional variance can easily be seen by comparing with the $\mathbb{A}_0(3)$ model. According to the latter model, S_t is equal to I in the diffusion term of the factor process $dX_t = \mathcal{K}(\Theta - X_t)\,dt + \sqrt{S_t}dB_t$. In contrast, the S-matrix in the $\mathbb{A}_1(3)$ model is given by

$$S_t = \begin{bmatrix} X_{1t} & 0 & 0 \\ 0 & 1+\beta_{12}X_{1t} & 0 \\ 0 & 0 & 1+\beta_{13}X_{1t} \end{bmatrix}. \tag{41}$$

The first factor, X_1, therefore enters as a stochastic volatility factor in each of the three factors. However, as argued above, this added flexibility in the variance specification comes at a cost in terms of the flexibility of the price of risk in the essentially affine framework. Specifically, because $\inf\left(\alpha_1 + \beta_1' X_t\right)$ is no longer greater than zero, (26) requires that $S^-_{t(11)} = 0$, resulting in

$$\sqrt{S_t^-} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & (1+\beta_{12}X_{1t})^{-1/2} & 0 \\ 0 & 0 & (1+\beta_{13}X_{1t})^{-1/2} \end{bmatrix}. \tag{42}$$

Since the first column of $\sqrt{S_t^-}$ is zero, the first row of ψ does not impact on the market price of risk and we therefore set this row to zero.

As in the $\mathbb{A}_0(3)$ case, we estimate the $\mathbb{A}_1(3)$ model using the Kalman filter technique. However, in contrast to the Gaussian case, there is no longer a closed-form solution available for the conditional density of bond yields, and we can therefore not use exact ML estimation. A number of alternative approaches have been suggested in the literature, including Simulated Method of Moments (Duffie and Singleton (1993)), Efficient Method of Moments (Gallant and Tauchen

(1996)), and estimation based on the conditional characteristic function (Singleton (2000)). However, these approaches often turn out to be quite computationally burdensome to implement, in particular for multi-factor ATSMs. We therefore proceed with the use of the Kalman filter and assume a normal distribution for the innovations in the transition equation in order to implement a quasi-maximum likelihood (QML) estimation approach, as in Lund (1997), De Jong (2000), and Duffee and Stanton (2000), among others. In this context, it is interesting to note that Duffee and Stanton (2000) find in a Monte Carlo study of the EMM and the Kalman filter/QML estimation techniques that the latter performed substantially better than the former in small samples.

Compared with the $\mathcal{A}_0(3)$ case, the implementation of the estimation procedure for the $\mathcal{A}_1(3)$ model requires a few adjustments. The measurement equation (27) of the state space model will remain unchanged, but the functions $A(\tau)$ and $B(\tau)$ will now be solutions to a different system of ODEs (not shown), reflecting the new dynamics and risk prices of the factors. The numerical Runge-Kutta method can be used to provide fast and accurate solutions. The transition equation, however, will be different because Θ no longer has all elements equal to zero. Specifically, we get

$$X_t = c(\phi) + \Phi(\phi) X_{t-h} + u_t, \tag{43}$$

where it can be shown that (see e.g., Lund (1997) or Duffee (2002))

$$\begin{aligned} c(\phi) &= \int_{t-h}^{t} \exp[-\mathcal{K}(t-s)]\, \mathcal{K}\Theta ds \\ &= (I - \exp[-\mathcal{K}(h)])\, \Theta, \end{aligned} \tag{44}$$

while $\Phi(\phi)$ is defined as in (34). Next, we assume that the innovations in the transition equation are multivariate normally distributed in order to implement the QML estimation procedure. Because the factor variances now are time-varying, the covariance matrix of the innovations is given by

$$V_t(\phi) = \int_{t-h}^{t} \exp[-\mathcal{K}(t-s)]\, S_t \exp\left[-\mathcal{K}'(t-s)\right] ds. \tag{45}$$

Parameter estimates, obtained as outlined above, are reported in Table III. Table IV presents a number of statistics for in-sample yield pricing errors (observed minus model yields) corresponding to this specification, as well as for the $\mathcal{A}_0(3)$ model to enable a comparison. Apart from the mean, standard deviation and first-order autocorrelation of the pricing errors, the table also displays the "Q-invert" and the "Q-steep" statistics used by Dai and Singleton (2000) to investigate the performance of various affine models. These statistics show the sample means of the pricing errors corresponding to the months when the slope of the yield curve (10-year − 1-year yields) was in the lowest quartile (inverted) and in the

Table III. QML parameter estimates for the essentially affine $\mathbb{A}_1$ (3) model.

The standard errors are based on the asymptotic variance-covariance matrix of White (1982). "–" denotes that the parameter is estimated to be on the boundary of the admissible parameter space. The estimates of the measurement error variances in $H(\phi)$ are not reported.

Parameter	Estimate	Standard error
δ_0	0.044	0.004
$\delta_1 \times 10^2$	0.036	0.015
$\delta_2 \times 10^2$	0.000	–
$\delta_3 \times 10^2$	0.304	0.091
θ_1	29.229	7.813
κ_{11}	0.039	0.005
κ_{21}	−0.169	0.111
κ_{22}	0.417	0.152
κ_{23}	−0.002	0.006
κ_{31}	0.000	–
κ_{32}	−0.727	0.424
κ_{33}	0.949	0.237
β_{12}	2.216	3.029
β_{13}	0.461	0.328
λ_1	0.003	0.003
λ_2	−0.093	0.100
λ_3	−1.646	1.050
ψ_{21}	0.033	0.069
ψ_{22}	−0.241	0.178
ψ_{23}	0.261	0.115
ψ_{31}	0.787	0.819
ψ_{32}	−0.159	0.125
ψ_{33}	0.212	0.158

highest quartile (steep), respectively, among the observed slopes in the sample. They are useful for detecting systematic patterns of mispricing related to available information in the observed yield curve. For both models, we see a tendency of underpricing when the yield curve is inverted, and a tendency of overpricing when the yield curve is steep. However, in most cases, these tendencies appear limited.

The overall picture is that both models seem to do roughly equally well in terms of pricing yields in-sample: the average pricing errors are very small, the standard deviations and correlations are reasonably low, and the aforementioned statistics based on the slope of the yield curve do not signal any major misspecifications. However, one caveat seems warranted. There are signs that the two models perform

Table IV. Moments of pricing errors

All figures refer to statistics for the yield pricing errors expressed in basis points. "Corr." is the first-order autocorrelation of the pricing errors, "Q-invert" and "Q-steep" show the sample means of the pricing errors corresponding to the months when the slope of the yield curve (10-year minus 1-year yields) was in the lowest quartile and in the highest quartile, respectively, among the observed values.

Maturity	Mean	Std.	Corr.	Q-invert	Q-steep
$\mathcal{A}_1(3)$ model					
3 months	−1.9	30.9	0.36	17.2	−12.4
6 months	2.4	26.1	0.19	10.7	−1.2
9 months	−0.2	25.2	0.22	6.3	−4.0
1 year	−1.3	25.7	0.27	4.2	−5.1
2 years	−0.8	26.9	0.31	3.5	−3.9
3 years	0.1	26.1	0.29	4.0	−2.4
5 years	−0.1	24.2	0.23	3.5	−1.4
7 years	−0.5	23.0	0.15	1.7	−0.3
10 years	−0.8	23.2	0.13	−0.8	3.4
$\mathcal{A}_0(3)$ model					
3 months	−0.5	29.9	0.32	17.0	−7.1
6 months	3.3	26.1	0.19	9.9	3.2
9 months	0.3	25.2	0.21	5.1	−0.5
1 year	−1.1	25.6	0.27	2.6	−2.2
2 years	−1.4	26.6	0.31	1.0	−2.5
3 years	−0.9	25.7	0.29	1.1	−1.5
5 years	−1.4	23.8	0.23	0.5	−0.5
7 years	−1.9	22.6	0.15	−1.0	0.5
10 years	−0.4	22.9	0.14	−2.9	4.2

somewhat less well at the very short end of the term structure, as indicated by the slightly larger standard deviations and autocorrelations of pricing errors for the 3-month rate. The "Q-invert" and "Q-steep" statistics also indicate that the performance of the models for the 3-month maturity is less satisfactory than for other maturities. This is consistent with empirical evidence for the U.S. presented by Dai and Singleton (2002), who argue that it may be necessary to include a fourth factor to fully capture the dynamics of the very short end of the term structure.

As a further test of the performance of the two $\mathcal{A}(3)$ models discussed above, Table V shows some statistics related to the (in-sample) forecast performance of the models. The table reports root mean squared errors (RMSE) of yields for four different maturities, and for four forecast horizons up to 12 months ahead. In addition to the results for the two affine models, RMSEs for a random walk benchmark are shown. As pointed out by Duffee (2002), completely affine models almost always fail to perform better than a random walk in terms of forecasting yields, whereas he shows that the essentially affine specification is more successful in this respect. Our results are in line with his findings: in none of the cases considered does the random walk obtain a lower RMSE than either the $\mathbb{A}_0(3)$ or the $\mathbb{A}_1(3)$ model. We also find that the $\mathbb{A}_1(3)$ model consistently performs at least as well as its Gaussian counterpart, although the differences are very small. Finally, in Table V we also report test statistics corresponding to a test of the null hypothesis that the forecast errors for all four maturities are uncorrelated with the slope of the yield curve (10-year minus 1-year yields) at the forecast date. This test (labelled "Moment test" in the table) is due to Duffee (2002). The idea is to test four overidentifying restrictions (one for each maturity considered), corresponding to the moments $E\left[\left(e^j_{t,t+T} - \overline{e^j_{t,t+T}}\right)(s_t - \overline{s_t})\right] = 0$, where $e^j_{t,t+T}$ is the forecast error, corresponding to the forecast made at t for the j-maturity yield at the future date $t+T$, and s_t is the slope of the yield curve observed at t. The test statistic is calculated along the lines of a GMM objective function, using an optimal Newey-West (1987) weighting matrix to account for overlapping observations in the moment conditions. The test statistic is distributed as χ^2 with 4 degrees of freedom under the null hypothesis. We find that both models considered do reasonably well in capturing the forecasting performance of the term-structure slope, although the Gaussian model seems to perform slightly better in general. This result is in line with the findings of Duffee (2002), who shows that the added flexibility of the market prices of risk in the Gaussian model is important to capture time variations in expected bond returns.

We now go back and examine the differences between the risk-neutral and the objective yield distributions for the $\mathbb{A}_1(3)$ model. Unfortunately, in contrast with the Gaussian case, these distributions are not known in closed form, as already mentioned. One alternative option would therefore be to approximate the conditional densities by the use of Monte Carlo techniques. However, a more exact and less computationally burdensome approach is available by relying on Fourier transform analysis, as suggested by Duffie, Pan, and Singleton (2000). From their results, one can find an expression for the conditional probability that the yield on a zero-coupon bond at some future date is greater than some specific value. To be precise, the probability - under some probability measure, say Q - at time t that, at the future time T, the yield of zero-coupon bond maturing at $T+\tau$ will be greater than z can be written

$$\begin{aligned}\Pr^Q_t\left(y_T(\tau) \geq z\right) &= \Pr^Q_t\left(-\frac{A(\tau)}{\tau} - \frac{B(\tau)'}{\tau}X_T \geq z\right)\\ &= \Pr^Q_t\left(-\frac{B(\tau)'}{\tau}X_T \geq z + \frac{A(\tau)}{\tau}\right).\end{aligned}$$

Because the conditional characteristic function of X_T is known in closed form, this probability can, given technical regularity conditions, be expressed as (see Appendix C in Duffie, Pan, and Singleton (2000))

Table V. In-sample yield forecast performance: RMSEs and moment test

All figures except those labelled "Moment test" are root mean squared errors (RMSE) for yield forecasts 3, 6, 9, or 12 months ahead (in sample forecasts), expressed in percentage points. "Moment test" refers to a test of the null hypothesis that the forecast errors for all four maturities are uncorrelated with the slope of the yield curve (10-year minus 1-year yields) at the forecast date. The test statistic for this test is shown for each of the four forecast horizons considered. The test is implemented as described in Duffee (2002, pp. 422–423), using a Newey-West weighting matrix with n+1 lags (where n is the number of months in the forecast horizon) . Figures in parentheses are p-values.

	Forecast horizon			
Maturity	3 months	6 months	9 months	12 months
$\mathcal{A}_1$ (3) model				
6 months	0.48	0.72	0.95	1.21
1 year	0.52	0.77	0.99	1.24
3 years	0.52	0.76	0.95	1.15
10 years	0.40	0.60	0.75	0.90
Moment test	3.98	9.77	9.27	8.68
	(0.409)	(0.045)	(0.055)	(0.070)
$\mathcal{A}_0$ (3) model				
6 months	0.48	0.73	0.96	1.23
1 year	0.52	0.77	0.99	1.25
3 years	0.53	0.77	0.96	1.17
10 years	0.40	0.61	0.76	0.91
Moment test	2.99	6.98	7.75	8.85
	(0.560)	(0.137)	(0.101)	(0.065)
Random Walk				
6 months	0.49	0.76	1.01	1.27
1 year	0.53	0.79	1.01	1.27
3 years	0.53	0.78	0.98	1.21
10 years	0.40	0.61	0.76	0.92

$$\Pr_t^Q \left(y_T(\tau) \geq z \right)$$

$$= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\mathrm{Im}\left[\psi^Q \left(-iv\frac{B(\tau)}{\tau}, X_t, t, T \right) \exp\left(-iv\left(z + \frac{A(\tau)}{\tau} \right) \right) \right]}{v} dv, \qquad (46)$$

where i is the imaginary unit $\sqrt{-1}$, and

$$\psi^Q \left(-iv\frac{B(\tau)}{\tau}, X_t, t, T \right) = E_t^Q \left[\exp\left(-iv\frac{B(\tau)}{\tau} X_T \right) \right] \qquad (47)$$

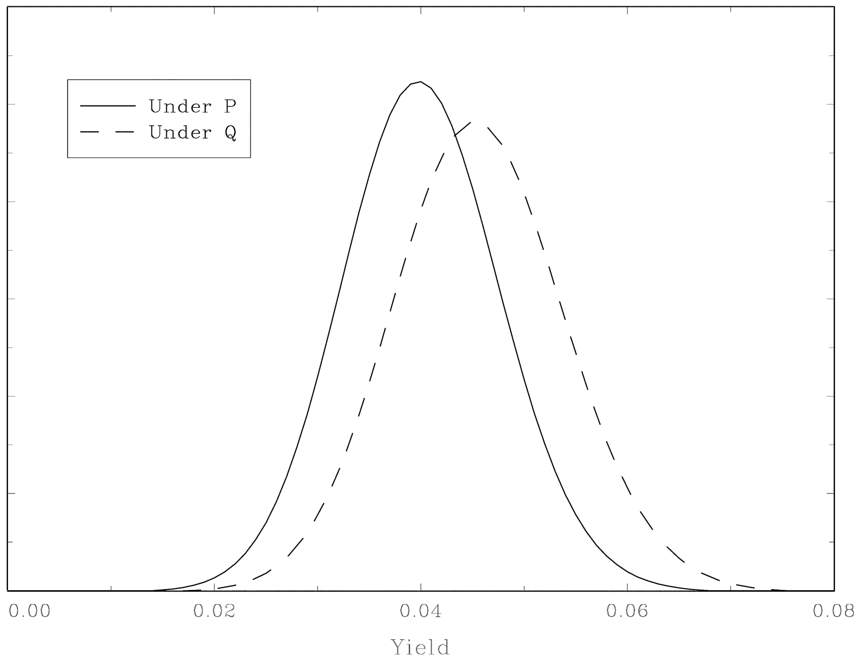

Figure 7. One-year ahead conditional distribution of the 3-month interest rate, as implied by the $\mathcal{A}_1$ (3)-model, end-March 2002.

denotes the conditional characteristic function. As shown by Duffie et al. (2000) and Singleton (2001), this characteristic function is for the affine case (under technical conditions) conveniently given by

$$\psi^{Q}\left(-iv\frac{B(\tau)}{\tau}, X_t, t, T\right) = \exp\left(\alpha(t) + \beta(t)\cdot X_t\right), \tag{48}$$

where $\alpha(t)$ and $\beta(t)$ satisfy a system of complex-valued ODEs with suitable boundary conditions (see also Appendix A in Hördahl and Vestin (2003)).

Figures 7 and 8 display the results obtained using the estimation procedure described above for the 3-month and 10-year zero-coupon yields at the one year horizon, as at end-March 2002. As in the case of the Gaussian specification, the results indicate that there are substantial differences between the risk-neutral and the objective yield densities. In the examples shown in Figures 7 and 8, the differences between the means of the one-year ahead Q and P distributions are 55 basis points for the 3-month rate, and 22 basis points for the 10-year yield. Moreover, also the standard deviations differ between the two probability measures, with the Q distributions having around 5-10% larger standard deviations than the P distributions.

Perhaps more important than these differences in the moments of the distributions, are the differences between the implied probabilities of specific outcomes,

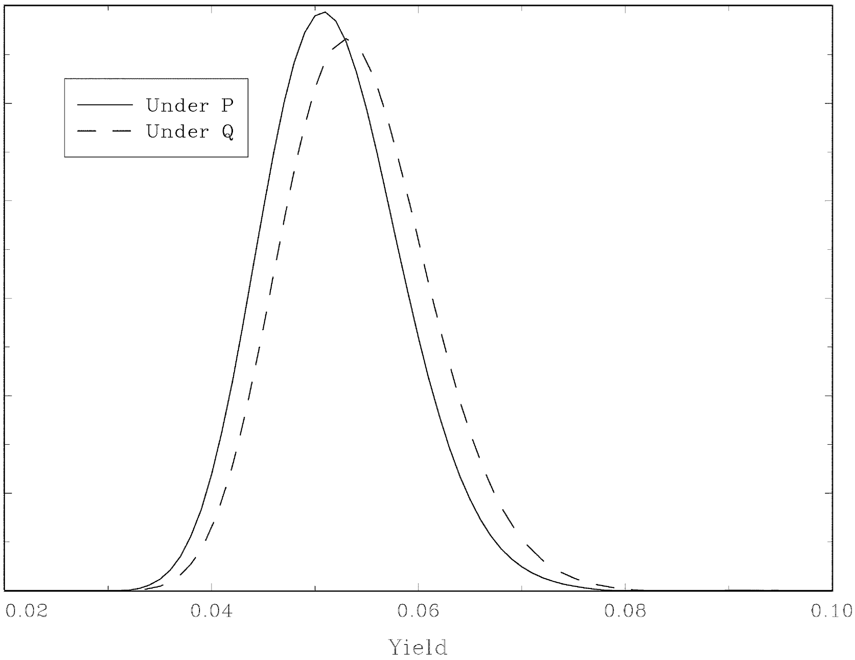

Figure 8. One-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the $\mathcal{A}_1$ (3)-model, end-March 2002.

since this type of information is typically what analysts focus on when interpreting densities implied by observed option prices. Given the estimates obtained for the $\mathcal{A}_1$ (3) model, our results show that these conditional probabilities can differ substantially between the two probability measures. For example, in the case of the one-year ahead 3-month rate in Figure 7, the risk-neutral probability of an outcome above 5.00% is around 25%, while the true objective probability for this outcome is only 8%. Also in the case of the one-year ahead 10-year yield we find large differences: in the example above, the Q-probability of an outcome above 5.50% is 34%, whereas the corresponding P-probability is 18%. These differences in the moments of the distributions and in the implied probabilities of future outcomes are by no means extreme. Table VI summarizes some of the statistics regarding the differences in the moments over the entire sample period, as estimated by the $\mathcal{A}_1$ (3) model.

Comparing with the results obtained from the $\mathcal{A}_0$ (3)-model, a few similarities and some differences are worth pointing out. First, focusing on the 10-year densities in Figures 6 and 8, the $\mathcal{A}_0$ (3) model implies that the risk-neutral density is centered somewhat to the left of the objective density, while the $\mathcal{A}_1$ (3) model produces the opposite result. A closer look at the relation between the estimated Q and P densities from the two models shows that the Gaussian model produces frequent shifts in the relative positioning of the two distributions, whereas for the

Table VI. Differences between Q and P means and standard deviations for one-year ahead 3-month and 10-year yields, as implied by the Essentially Affine $\mathcal{A}_1(3)$ model, January 1983–March 2002

Mean figures refer to the level difference $(Q - P)$ in basis points; standard deviation figures refer to the relative difference in percent of the P-distribution volatility.

	Average	Minimum	Maximum
Mean 3-m. rate	53 bps	11 bps	105 bps
Mean 10-yr. yield	25 bps	9 bps	42 bps
St. dev. 3-m. rate	8.4%	6.0%	8.5%
St. dev. 10-yr. yield	4.6%	2.2%	4.7%

$\mathcal{A}_1(3)$ model, the risk-neutral density is always centered to the right of the objective density (see Figures 9 and 10). While we have not examined the specific reasons for these discrepancies, we conjecture that it is the imposed constant-variance assumption in the Gaussian model that drives the swings in the relationship between the two distributions.

Second, in the example above which relates to end-March 2002, we see that the densities are more dispersed in the $\mathcal{A}_0(3)$ case than in the $\mathcal{A}_1(3)$ case. This is due to the fact that while the yield variance appeared to be relatively low in early 2002, the former model imposes a constant yield variance for any given maturity. Hence, in order to get the average variance right, the $\mathcal{A}_0(3)$ model overstates the variance in times of relatively low volatility, whereas the $\mathcal{A}_1(3)$ model is better equipped to capture swings in the volatility. The estimated time series of $\mathcal{A}_1(3)$-variances for the one-year ahead densities, displayed in Figures 11 and 12 for the 3-month and the 10-year cases respectively, show a significant degree of variability over time. This would tend to suggest that stochastic volatility should be an important feature for term structure models aimed at capturing salient features of actual observed data.

Third, in both cases, the dispersion of the 3-month PDFs is greater than for the 10-year counterparts. This is due to the typically downward-sloping term structure of yield volatilities observed in actual yield curve data. If a term structure model is successful in capturing this feature, one would expect to see a higher variance in short-rate densities than in long-term yield densities.

As discussed earlier, the flexibility of the$\mathcal{A}_1(3)$ model with respect to capturing time-varying volatility has a cost in terms of reduced flexibility in the mean specification. As Duffee (2002) shows, the specification of the mean is important in order to get the expected excess return right. Roughly, this excess return is captured by the difference in mean of the P and Q distributions. Empirically, the observed excess return is a volatile component that switches sign over time, and hence the $\mathcal{A}_0(3)$ model utilises its flexible risk premia specification to match this feature of the data more closely. Ultimately, which model is to be prefered must be guided by the aim in mind. Since Duffee (2002) focuses on conditional mean forecasts, he ends up favoring the $\mathcal{A}_0\,(3)$ model. In this paper, however, we are interested in the models' ability to capture the entire distribution of future yields, in which case it seems likely that also time-varying volatility will be important. The choice of model should therefore be guided by an evaluation of the density forecast accuracy, which is the aim of the next section.

Leaving aside the results from the Gaussian model and focusing on the $\mathcal{A}_1\,(3)$ results, one can conclude that there seem to be relatively substantial differences between estimated risk-neutral and objective yield densities at the one-year horizon. As could be expected, the differences between the risk-neutral and the objective densities tend to decline as the horizon is reduced. This can be seen, for example, by comparing the plots of the means of the one-year ahead densities from the $\mathcal{A}_1\,(3)$ model in Figure 9 and 10 with the corresponding 6-month and 3-month ahead PDFs displayed in Figures 13 and 14, respectively. At times, however, non-negligible differences remain even for relatively short horizons. Moreover, the results show that not only the differences between the means of the distributions can vary substantially over time, but that there also seems to be considerable time-variation in the estimated variances of the Q and P densities. These findings therefore point to the need for a high degree of caution when interpreting risk-neutral densities such as option-implied PDFs in terms of actual objective probabilities.

5. Density Evaluation

The previous section presented estimates of both risk-neutral and objective interest rate PDFs, and provided some illustrative evidence that these densities differed in important ways, as well as in a time-varying manner. A natural question to ask at this point is how reasonable these estimates are. This section aims at shedding some light on this question. In this context, it is intresting to note that Egorov, Hong and Li (2003) extends the analysis and methodology of Duffee (2002) to include a density forecast evaluation on U.S. data. They find evidence that the $\mathcal{A}_1\,(3)$ model is best suited to capture the features of the U.S. term structure, hence providing evidence from an independent data set that the restrictions implied by this model appropriately balances the trade-off in terms of the flexibility in the mean and the volatility of yields.

Starting with the question whether the risk-neutral densities have been estimated reasonably well, we merely conclude that the estimated Q-densities are consistent with the bond-pricing equations used in the estimation process, and that these in turn do a good job in terms of pricing the bonds (see also Table IV). An alternative way to evaluate the accuracy of the risk-neutral densities would be to examine how well they would be able to price traded interest rate derivatives. We leave this for future research.

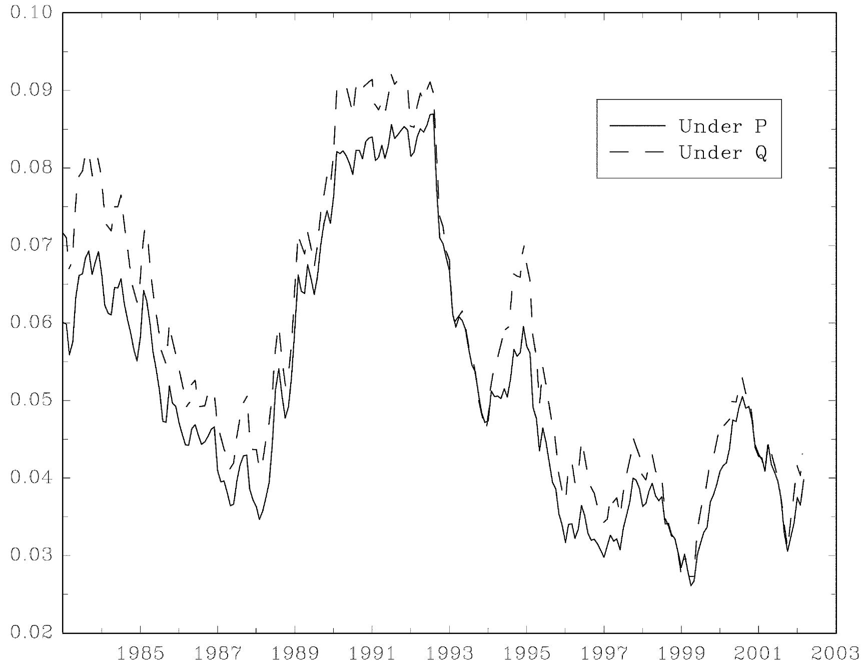

Figure 9. The mean of the one-year ahead conditional distribution of the 3-month interest rate, as implied by the $\mathcal{A}_1$ (3)-model, January 1983–March 2002.

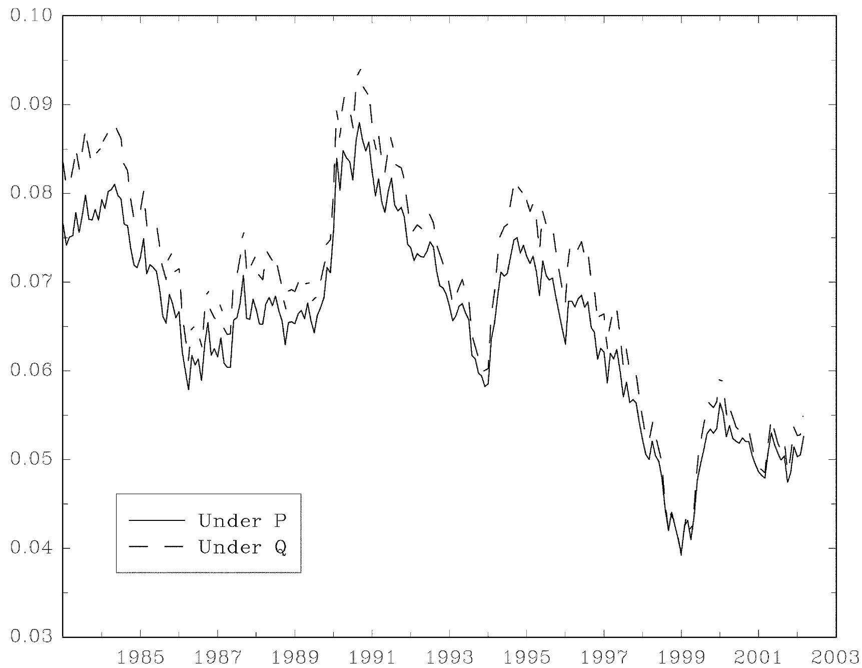

Figure 10. The mean of the one-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the $\mathcal{A}_1$ (3)-model, January 1983–March 2002.

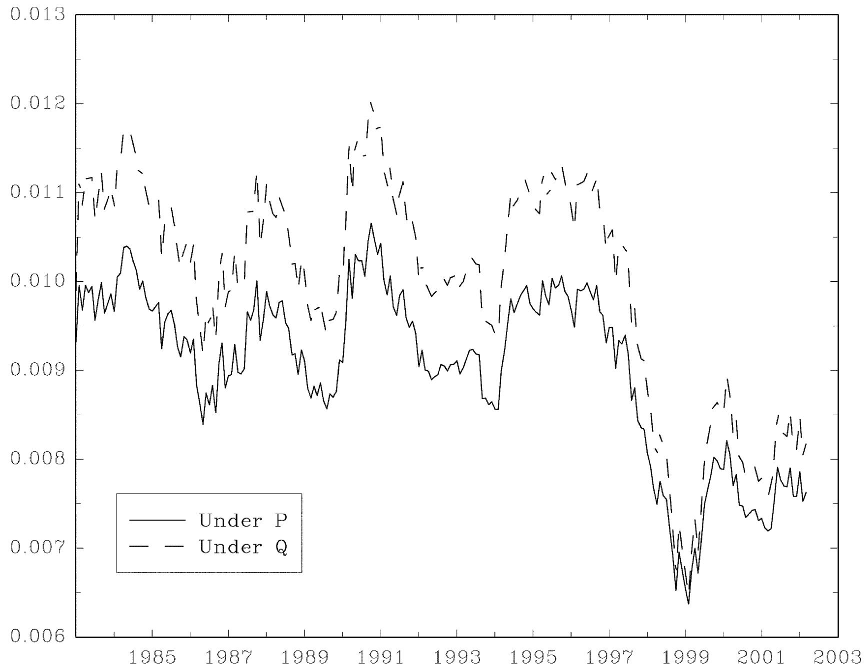

Figure 11. The square-root of the variance of the one-year ahead conditional distribution of the 3-month interest rate, as implied by the $\mathcal{A}_1(3)$-model, January 1983–March 2002.

Turning to the accuracy of the estimated physical densities, we employ a recently proposed density forecast evaluation methodology to investigate this issue. Tests of this kind can help us determine whether the estimated objective interest rate densities for any given forecast horizon work well in terms of describing the distribution of realized interest rates at that horizon. Building on the results of Rosenblatt (1952), Diebold et al. (1998) proposed a method for evaluating a sequence of density forecasts using a probability integral transform of the realized outcomes of the forecasted variable. Specifically, denoting the outcome r_t and the ex-ante density forecast $f_t(\cdot)$, the proposed transformation is defined as

$$\begin{aligned} x_t &= \int_{-\infty}^{r_t} f_t(u)\, du \\ &= F_t(r_t). \end{aligned} \tag{49}$$

Diebold et al. (1998) show that if a sequence of density forecasts, $\{f_t(r_t)\}_{t=1}^m$ coincides with the true density sequence, then the sequence of probability integral transforms of $\{r_t\}_{t=1}^m$ with respect to $\{f_t(r_t)\}_{t=1}^m$ is IID with uniform distribution,

$$\{x_t\}_{t=1}^m \overset{\text{IID}}{\sim} U(0,1).$$

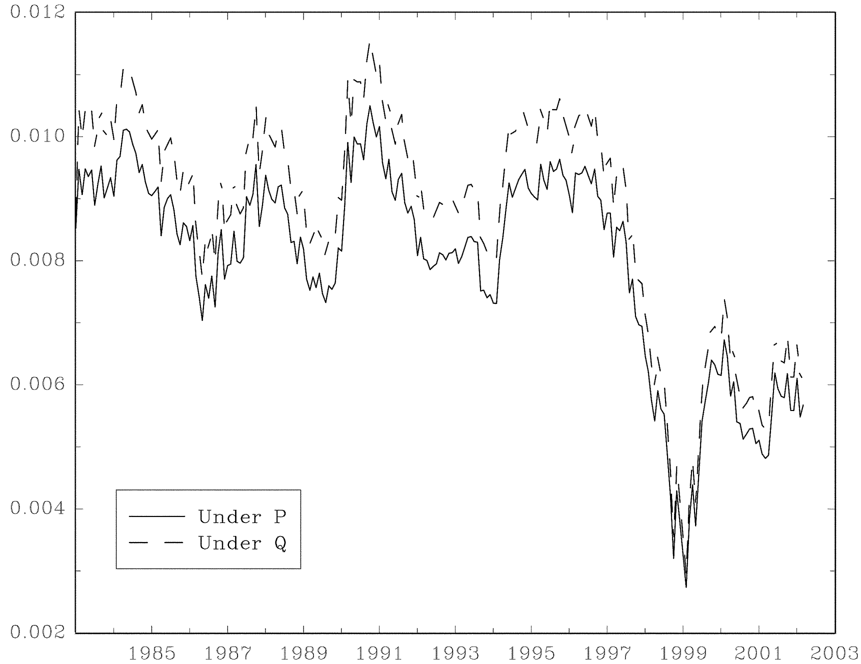

Figure 12. The square-root of the variance of the one-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the $\mathcal{A}_1(3)$-model, January 1983–March 2002.

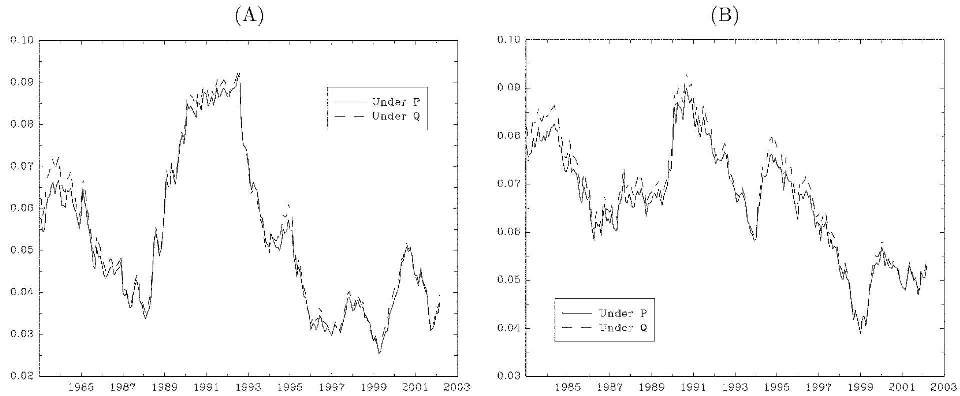

Figure 13. The mean of the 6-month ahead conditional distribution of the 3-month interest rate (panel A) and of the 10-year zero-coupon bond yield (panel B), as implied by the $\mathcal{A}_1(3)$-model, January 1983–March 2002.

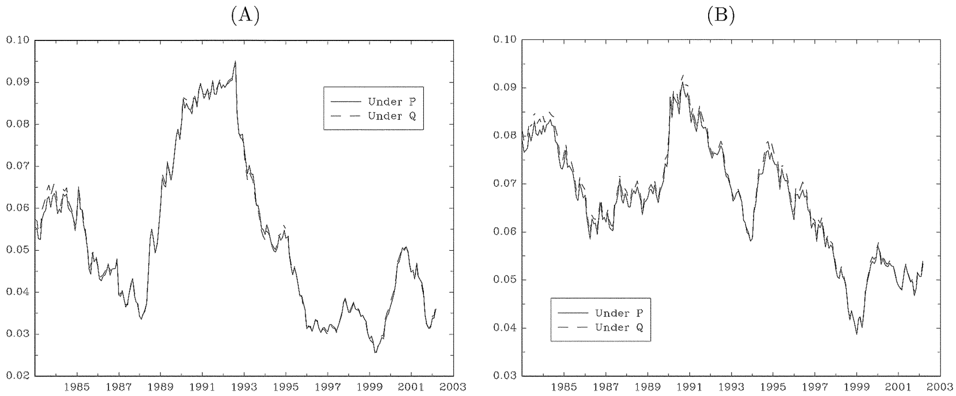

Figure 14. The mean of the 3-month ahead conditional distribution of the 3-month interest rate (panel A) and of the 10-year zero-coupon bond yield (panel B), as implied by the $\mathcal{A}_1$ (3)-model, January 1983–March 2002.

Diebold et al. (1998) did not focus on formal testing procedures, but instead suggested the use of simple visual tools such as histograms and correlograms to ascertain whether the transformed series x is IID U (0, 1).

More recently, Berkowitz (2001) extended the Diebold et al. (1998) approach to allow for formal testing of the density forecasts. This new approach relies on a simple transformation of the x-series in (5) into a normal distribution (under the null). More precisely, if x_t is IID U (0, 1), then the transformed series

$$\begin{aligned} z_t &= \Phi^{-1}\left[\int_{-\infty}^{r_t} f_t(u)\,du\right] \\ &= \Phi^{-1}(x_t) \end{aligned} \tag{50}$$

will be IID N (0, 1), which, as Berkowitz (2001) notes, can be tested using e.g., LR tests. We implement this testing methodology for objective densities obtained from the estimated $\mathcal{A}_0$ (3) and $\mathcal{A}_1$ (3) models. Specifically, we consider the same maturities discussed in previous sections (3-month and 10-year rates) and we focus on two forecast horizons: 3 months ahead and 12 months ahead. For each combination of model, horizon and maturity, the null hypothesis that the transformed series z is normally distributed with zero mean, unit variance, and zero first-order autocorrelation ($\mu_z = 0, \sigma_z = 1, \rho_z = 0$) is tested against four different alternatives as follows:

Test	$H_{alt.}$
1	$\mu_z = \hat{\mu}_z, \sigma_z = 1, \rho_z = 0$
2	$\mu_z = 0, \sigma_z = \hat{\sigma}_z, \rho_z = 0$
3	$\mu_z = \hat{\mu}_z, \sigma_z = \hat{\sigma}_z, \rho_z = 0$
4	$\mu_z = \hat{\mu}_z, \sigma_z = \hat{\sigma}_z, \rho_z = \hat{\rho}_z$

We implement the tests by obtaining the log-likelihood values under the null hypothesis and the various alternatives, and then calculate the corresponding likelihood ratio statistics. All tests are in-sample tests, in the sense that the density forecasts at a given forecast date are conditional on the information available about the term structure and the state variables at that time, given the parameter values for the model obtained using the entire sample.[10]

Table VII displays the results for both the $\mathcal{A}_0(3)$ and $\mathcal{A}_1(3)$ models. Starting with the Gaussian model in Panel A, the test results show that this model does not perform very well at the 12-month forecast horizon: in 7 out of 8 cases, the null hypothesis can be rejected. In other words, objective densities estimated with the $\mathcal{A}_0(3)$ model do not seem to coincide with the true densities at the 12-month horizon. For the 3-month ahead forecast horizon, the results are more mixed, since we we reject in half of the cases, for both maturities. The critical element in these rejections seems to be that μ_z differs significantly from zero, meaning that the conditional means of the estimated densities differ systematically from the realizations.

Next, we turn to Panel B in Table VII, which shows the results for the $\mathcal{A}_1(3)$ model. Starting again with the 12-month forecast horizon, we find that the $\mathcal{A}_1(3)$ model does considerably better than the Gaussian model. In 6 out of 8 cases we cannot reject the null hypothesis for the estimated objective densities. In particular, for the 10-year bond yield densities, the null is not rejected in any of the four cases. The two rejections appear to be due to a failure in capturing the conditional volatility of the 3-month interest rate. For the 3-month forecast horizon, the estimated objective densities again seem to do a relatively good job in terms of capturing the behavior of the outcomes. As is the case of the longer forecast horizon, we fail to reject the null in 6 out of 8 cases. This time the rejections seem to be due to correlated forecast errors for the 3-month rate, and to differences between the modelled and the actual volatility of the 10-year yield at the 3-month horizon. It is worthwhile pointing out that in none of the cases do we obtain any rejection as a result of significant deviations of μ_z from zero. This indicate that the $\mathcal{A}_1(3)$ model seems to do a good job in terms of capturing the time-varying level of risk premia at both horizons and for both maturities. Overall, we conclude from these tests that the $\mathcal{A}_1(3)$ model is able to capture objective forward-looking densities in a much more satisfactory way than the Gaussian model, although some room for improvement still remains.[11]

[10] Ideally, an out-of-sample testing approach would have been preferable. However, the limited sample period and the need for a relatively large number of realizations to obtain meaningful test results meant that an out-of-sample approach had to be ruled out in practice.

[11] We also performed similar test on the estimated Q-densities' ability to capture the distribution of the realizations (results not shown). As could be expected, the risk-neutral densities did not perform very well in this regard, in particular at the one-year ahead horizon where we rejected the null in all 16 cases. This again highlights the dangers associated with relying on RNDs as measures of expectations.

Table VII. Berkowitz (2001) density forecast evaluation tests

The test statistics are chi-squared distibuted with 1 degree of freedom for tests 1 and 2, 2 degrees of freedom for test 3, and 3 degrees of freedom for test 4. * denotes statistical significance at the 5% level. Monthly density forecasts are used in all tests, except in test 4, where non-overlapping observations are used, since these tests include the autocorrelation coefficient ρ.

PANEL A: Evaluation of $\mathcal{A}_0$ (3) densities				
	3-month forecast horizon			
	LR_1	LR_2	LR_3	LR_4
3m. rate/P-measure	13.24*	0.13	12.93*	7.31
10y. rate/P-measure	4.38*	3.30	7.06*	6.89
	12-month forecast horizon			
	LR_1	LR_2	LR_3	LR_4
3m. rate/P-measure	4.81*	38.76*	41.62*	7.27
10y. rate/P-measure	15.79*	26.12*	36.69*	11.43*
PANEL B: Evaluation of $\mathcal{A}_1$ (3) densities				
	3-month forecast horizon			
	LR_1	LR_2	LR_3	LR_4
3m. rate/P-measure	1.66	1.03	2.86	8.21*
10y. rate/P-measure	0.03	4.19*	4.22	4.38
	12-month forecast horizon			
	LR_1	LR_2	LR_3	LR_4
3m. rate/P-measure	1.80	25.49*	26.65*	6.57
10y. rate/P-measure	3.82	1.49	4.91	6.97

6. Conclusions

A commonly used approach to extract information on market expectations and the perceived degree of uncertainty about some asset price is to estimate the implied density using observed prices of options written on the asset of interest. However, it is well known that this approach produces the risk-neutral density, rather than the true objective density. This paper examines the differences between risk-neutral and objective probability densities for interest rates and bond yields. This kind of analysis may be of interest for e.g. central banks, since RNDs are commonly employed by central banks as indicators of market expectations for underlying macroeconomic fundamentals as well as future monetary policy. Instead of modeling the terminal distribution directly, the approach taken here is to model the dynamics of the underlying state variables. Specifically, a multi-factor essentially affine modeling framework is applied to German time-series and cross-section term structure data in order to identify both the risk-neutral and the objective term structure dynamics.

In general, the results show that there are important differences between risk-neutral and objective distributions as a result of risk premia in bond prices. For example, the one-year ahead distributions for the three-month interest rate and the ten-year bond yield display substantial differences in both the means and the variances of the two types of distributions. While the magnitude of these differences diminish as the forecast horizon is shortened, important differences remain for horizons that are commonly used in practice. The results also indicate that the differences between the objective and the risk-neutral distributions vary over time, as a result of time-varying risk premia. Moreover, density forecasts performed on estimated objective densities show that the proposed approach does reasonably well in terms of capturing the true realized densities (for the $\mathcal{A}_1(3)$ model). We therefore conclude that one should be cautious in interpreting RNDs, such as option-implied densities, in terms of expectations. The method used in this paper provides one alternative approach which can be used to identify risk premia and thereby the objective probabilities of future outcomes.

Appendix

Given the state-space specification for the $\mathcal{A}_0(3)$ model, the prediction step follows from the transition equation, resulting in (suppressing dependence on ϕ)

$$\begin{aligned}\hat{Y}_{t|t-h} &= E\left[Y_t \mid R_{t-h}\right] \\ &= \Phi \hat{Y}_{t-h},\end{aligned} \tag{51}$$

and an associated MSE matrix,

$$\begin{aligned}\hat{\Omega}_{t|t-h} &= E\left[\left(Y_t - \hat{Y}_{t|t-h}\right)\left(Y_t - \hat{Y}_{t|t-h}\right)' \mid R_{t-h}\right] \\ &= \Phi \hat{\Omega}_{t-h} \Phi' + V_t.\end{aligned} \tag{52}$$

Next, the observed yields R_t are used in the update step, to provide the filtered estimator of Y :

$$\begin{aligned}\hat{Y}_t &= E\left[Y_t \mid R_t\right] \\ &= \hat{Y}_{t|t-h} + \hat{\Omega}_{t|t-h} Z' F_t^{-1} \upsilon_t,\end{aligned} \tag{53}$$

where υ_t is the vector of prediction errors,

$$\upsilon_t = R_t - \left(d + Z\hat{Y}_{t|t-h}\right), \tag{54}$$

and F_t is the covariance matrix of the prediction errors,

$$\begin{aligned} F_t &= cov\left(R_t \mid R_{t-h}\right) = E\left[\upsilon_t \upsilon_t' \mid R_{t-h}\right] \\ &= Z\hat{\Omega}_{t|t-h}Z' + H. \end{aligned} \tag{55}$$

The update of the MSE matrix Ω is given by

$$\begin{aligned} \hat{\Omega}_t &= E\left[\left(Y_t - \hat{Y}_t\right)\left(Y_t - \hat{Y}_t\right)' \mid R_t\right] \\ &= \hat{\Omega}_{t|t-h} - \hat{\Omega}_{t|t-h}Z'F_t^{-1}Z\hat{\Omega}_{t|t-h} \\ &= \left(I - \hat{\Omega}_{t|t-h}Z'F_t^{-1}Z\right)\hat{\Omega}_{t|t-h}. \end{aligned} \tag{56}$$

The prediction errors and their covariance matrix serves as input into the log-likelihood function of the Gaussian state-space model, resulting in the following likelihood function:

$$\begin{aligned} \ln L &= \sum_{t=1}^{T} \ln L_t, \\ \ln L_t &= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln|F_t| - \frac{1}{2}\upsilon_t' F_t^{-1}\upsilon_t. \end{aligned} \tag{57}$$

References

Aït-Sahalia, Y. (1996) Nonparametric pricing of interest rate derivative securities, *Econometrica* **64**, 527–560.

Aït-Sahalia, Y., Wang, Y., and Yared, F. (2001) Do option markets correctly price the probabilities of movement of the underlying asset?, *Journal of Econometrics* **102**, 67–110.

Björk, T. (1998) *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford.

Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities, *Journal of Political Economy* **81**, 659–683.

Bliss, R. R. and Panigirtzoglou, N. (2004) Option-implied risk aversion estimates, *Journal of Finance* **59**, 407–446.

Chacko, G. (1999) Continous-time estimation of exponential separable term structure models: A general approach, manuscript, Harvard University.

Cox, J., Ingersoll, J., and Ross, S. (1985) A theory of the term structure of interest rates, *Econometrica* **53**, 385–408.

Dai, Q. and Singleton, K. (2000) Specification analysis of affine term structure models, *Journal of Finance* **55**, 1943–1978.

Dai, Q. and Singleton, K. (2002) Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* **63**, 415–441.

de Jong, F. (2000) Time series and cross-section information in affine term-structure models, *Journal of Business & Economic Statistics* **18**, 300–314.

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

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

Duffee, G. and Stanton, R. (2001) Estimation of dynamic term structure models, Working paper, U.C. Berkeley.

Duffie, D. (2001) *Dynamic Asset Pricing Theory*, 3rd edn, Princeton University Press, Princeton, N.J.

Duffie, D. and Kan, R. (1996) A yield-factor model of interest rates, *Mathematical Finance* **6**, 379–406.

Duffie, D. and Singleton, K. (1993) Simulated moments estimation of Markov models of asset prices, *Econometrica* **61**, 929–952.

Duffie, D., Pan, J., and Singleton, K. (2000) Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**, 1343–1376.

Egorov, A., Hong, Y., and Li. H. (2003) Forecasting the joint probability density of bond yields: Can affine models beat the random walk?, Cornell University working paper (EFA 2003 meetings).

Gallant, A. R. and Tauchen, G. E. (1996) Which moments to match?, *Econometric Theory* **12**, 657–681.

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

Hördahl, P. and Vestin, D. (2003) Interpreting implied risk-neutral densities: The role of risk premia, Working Paper No. 274, European Central Bank, Frankfurt.

Lund, J. (1997) Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates, Working paper, Aarhus School of Business.

Melick, W. R. and Thomas, C. P. (1997) Recovering an asset's implied PDF from option prices: An application to crude oil during the Gulf crisis, *Journal of Financial and Quantitative Analysis* **32**, 91–115.

Nelson, C. R. and Siegel, A. F. (1987) Parsimonious modeling of yield curves, *Journal of Business* **60**, 473–489.

Newey, W. K. and West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

Singleton, K. (2001) Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* **102**, 111–141.

Söderlind, P. and Svensson, L. E. O. (1997) New techniques to extract market expectations from financial instruments, *Journal of Monetary Economics* **40**, 383–430.

Svensson, L. E. O. (1994) Estimating and interpreting forward interest rates: Sweden 1992–94, IMF Working paper WP/94/114, Washington D.C.

Vasicek, O. (1977) An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**, 177–188.

White, H. (1982) Maximum likelihood estimation of misspecified models, *Econometrica* **50**, 1–25.