

Hördahl, Peter; Vestin, David

**Working Paper**

## Interpreting implied risk-neutral densities: the role of risk premia

ECB Working Paper, No. 274

**Provided in Cooperation with:**

European Central Bank (ECB)

*Suggested Citation:* Hördahl, Peter; Vestin, David (2003) : Interpreting implied risk-neutral densities: the role of risk premia, ECB Working Paper, No. 274, European Central Bank (ECB), Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/152708>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

EUROPEAN CENTRAL BANK  
WORKING PAPER SERIES



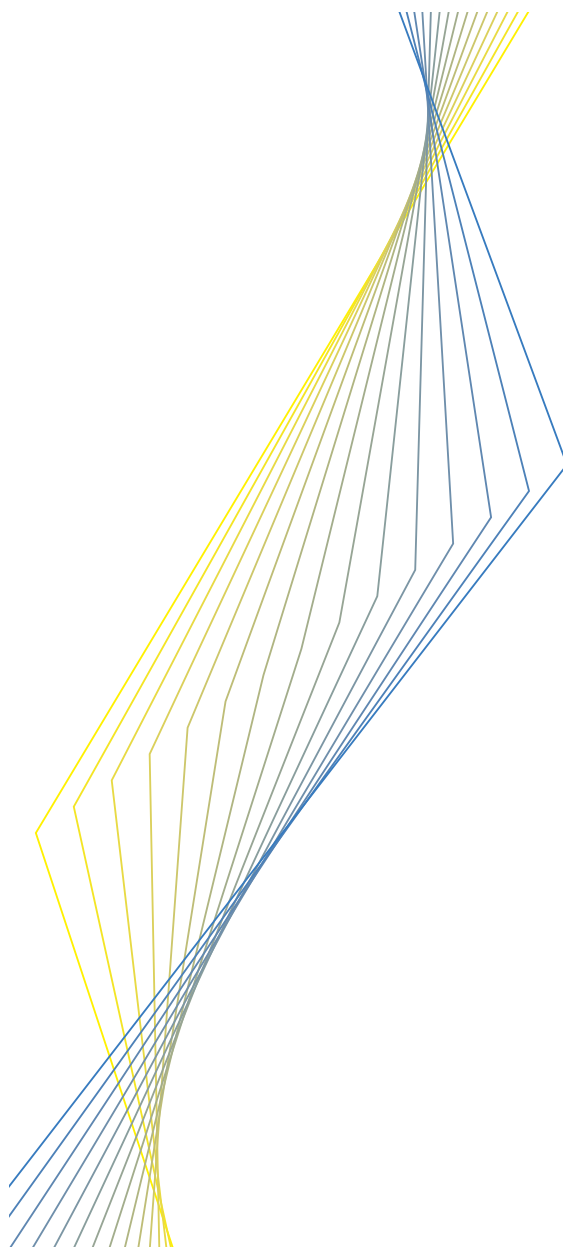
**WORKING PAPER NO. 274**

**INTERPRETING IMPLIED  
RISK-NEUTRAL DENSITIES:  
THE ROLE OF RISK PREMIA**

**PETER HÖRDAHL,  
DAVID VESTIN**

**SEPTEMBER 2003**

**EUROPEAN CENTRAL BANK**  
**WORKING PAPER SERIES**



**WORKING PAPER NO. 274**  
**INTERPRETING IMPLIED  
RISK-NEUTRAL DENSITIES:  
THE ROLE OF RISK PREMIA<sup>1</sup>**

**PETER HÖRDAHL<sup>2</sup>,  
DAVID VESTIN<sup>2</sup>**

**SEPTEMBER 2003**

<sup>1</sup> We are grateful for useful suggestions from Greg Duffee, Nikolaos Panigirtzoglou, Neil Pearson, Ranjini Sivakumar, 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 and the Oesterreichische Nationalbank. 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. This paper can be downloaded without charge from <http://www.ecb.int> or from the Social Science Research Network electronic library at [http://ssrn.com/abstract\\_id=487410](http://ssrn.com/abstract_id=487410).

<sup>2</sup> ECB, DG-Research, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany, tel.: +49-69-1344 6414 and 8691, e-mail: [peter.hordahl@ecb.int](mailto:peter.hordahl@ecb.int) and [david.vestin@ecb.int](mailto:david.vestin@ecb.int) respectively.

© European Central Bank, 2003

<b>Address</b>	<b>Kaiserstrasse 29 D-60311 Frankfurt am Main Germany</b>
<b>Postal address</b>	<b>Postfach 16 03 19 D-60066 Frankfurt am Main Germany</b>
<b>Telephone</b>	<b>+49 69 1344 0</b>
<b>Internet</b>	<b><a href="http://www.ecb.int">http://www.ecb.int</a></b>
<b>Fax</b>	<b>+49 69 1344 6000</b>
<b>Telex</b>	<b>411 144 ecb d</b>

*All rights reserved.*

*Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.*

*The views expressed in this paper do not necessarily reflect those of the European Central Bank.*

ISSN 1561-0810 (print)

ISSN 1725-2806 (online)

## Table of Contents

Abstract	4
Non-technical summary	5
1 Introduction	7
2 Theory	9
2.1 Modelling the distribution	10
2.2 Modelling the term structure	11
3 The Cox, Ingersoll and Ross Model	14
3.1 The CIR interest rate process	14
3.2 Risk premia in the CIR model	14
4 Three-factor essentially affine models	19
4.1 A gaussian three-factor ATSM: A0 (3)	21
4.2 Estimation using the kalman filter	24
4.3 The A1 (3) essentially affine model	29
5 Density evaluation	38
6 Conclusions	42
Appendices	
A Derivation of the ODEs for transform analysis	43
B The link between r- and P-distributions in a one-factor setting	44
C The bond price process in CIR	46
D Risk-neutral dynamics of $Y_t$ in the A0 (3) case	47
E The Kalman filter in the A0 (3) case	47
References	49
European Central Bank working paper series	51

## 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 conclude that one should be cautious in interpreting RNDs as representing the true expectations of market participants. The method used in this paper provides an alternative approach to identifying objective probabilities of future interest rates.

**JEL classification:** G12, E43

**Keywords:** Risk-neutral densities, essentially affine term structure models, risk premia

## Non-technical summary

Prices of financial contracts, such as bonds and futures, are commonly used to extract information about market expectations of future asset prices. This information, in turn, is often interpreted in terms of market participants' expectations of monetary policy and future underlying fundamentals, such as growth prospects and inflation. Central banks, for example, frequently use measures implicit in financial asset prices to obtain estimates of market expectations. A large number of techniques have been developed towards extracting information of this type, such as the modelling of implied forward rates. Another approach that has been increasingly used to examine measures of uncertainty, is to model the entire implied probability density of the price of a financial contract for a future date of interest. Prices of options and other derivatives are frequently used to derive such implied densities. This approach has the advantage of providing information not only about expectations of the level of future variables, but also about the degree of uncertainty associated with these expectations, as well as any possible asymmetries in terms of the range of expected future outcomes.

A well-known problem with the extraction of information about expectations of future asset price movements from market prices is that the existence of risk premia drives a wedge between discounted expected future asset prices and currently observable asset prices. For example, the current futures price of, say, an equity index in three months time is in general not equal to the market's expectation of that index three months ahead, because of the impact of risk premia.

Despite the fact that it is well-known that risk premia have this "distorting" impact, the specific role of premia is often explicitly or implicitly disregarded when expectations about future asset prices are extracted from observed prices. One example is the relatively common practice to interpret implied forward rates as direct measures of expected future interest rates. This paper examines the role of risk premia when examining and interpreting the information from financial indicators in the context of interest rate-related markets. In particular, the paper focuses on the impact of risk premia on implied probability densities of future interest rates.

Almost all available methods used to estimate implied densities deliver the so-called "risk-neutral density" of the underlying price or interest rate at the time of expiration, rather than the true "real-world" probability density. The reason why it is called the "risk-neutral density" is that it would correspond to the real-world density only if investors were in fact risk neutral. Again, the difference arises because of the existence of risk premia.

The strength of this approach, where risk-neutral densities are modelled directly, is that it avoids making explicit assumptions regarding the dynamic evolution of interest rates. However, the weakness is that there is little scope to assess the risk premia issue within that framework. To address this issue, we model the stochastic behavior of the term structure of interest rates by linking it to a number of underlying factors with certain properties. Once the dynamics of the underlying factors are specified, then we can derive the dynamics of interest rates for any maturity using no-arbitrage principles, and the distributional properties can then be uniquely pinned down. Moreover, given an assumption about the functional form of the risk premium in terms of the factors, it is possible to transform risk-neutral densities into real-world densities, and vice versa, once the parameters of the model have been estimated using term structure data. This allows us to quantify the differences between these two types of probability densities. For example, we can explore the differences in the means and the variances of the two densities for any combination of maturity and horizon, and examine if and how these differences vary over time.

The empirical analysis is carried out on German term structure data, and the interpretation of the results is focused on densities for 3-month interest rates and 10-year yields for horizons of up to one year. The results we obtain indicate that risk premia considerations are important, in the sense that there generally are non-negligible differences between risk-neutral and real-world densities. More importantly, we find that the differences between these densities change over time as a result of time-variation in risk premia. We therefore conclude that caution should be exercised when interpreting risk neutral densities, such as option-implied densities, as representing the true expectations of market participants. In addition, in an evaluation of the estimated real-world densities with respect to actual future outcomes, we find that we do a good job in terms of capturing the true densities of future interest rates. This indicates that the dynamics of risk premia are modelled satisfactorily in the framework we use. The method used in this paper therefore provides one alternative approach which can be used to identify risk premia and thereby the real-world probabilities of future outcomes.



# 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 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, the derivative pricing framework largely relies 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 (2001). 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. Specifically, 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.

Obviously, once the dynamics are specified, the distributional properties 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 6 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,  $\Pi$  at time  $t$ , based on an underlying asset  $Y$  at time  $T$ , will be given by

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

where  $\phi(Y_T)$  is the payoff of the claim and  $r(s)$  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 in 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 whole analysis is conducted under the risk-neutral measure.

To evaluate the quantitative differences, we therefore introduce a second approach, i.e. the modelling of the evolution of underlying state variables which 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,

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

## 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

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

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

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

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,  $\gamma = [\mu_1, \mu_2, \sigma_1, \sigma_2, \theta]'$ . Specifically, the call prices are

given by

$$\begin{aligned}
\tilde{C}(t, T; K, \mu, \sigma) &= E_t^Q [P(T) - K]^+ \\
&= \int_K^\infty (P(T) - K) f(P(T)) dP(T) \\
&= \theta e^{\mu_1 + \frac{\sigma_1^2}{2}} N(\sigma_1 - d_1) + (1 - \theta) e^{\mu_2 + \frac{\sigma_2^2}{2}} N(d_2 - \sigma_2) \\
&\quad - K (\theta N(-d_1) + (1 - \theta) N(-d_2)),
\end{aligned}$$

where

$$d_i = \frac{\log(K) - \mu_i}{\sigma_i}, \quad i = 1, 2.$$

Next, 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<sup>1</sup>

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

Further, if one is prepared to assume that the  $Q$ -measure coincides (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.<sup>2</sup> 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

The vast majority of the pricing literature models the dynamic evolution of the 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 parts, together with initial conditions for the state variables

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

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. It

<sup>1</sup>The purist may note that if the assumed density shape is correct, the resulting errors will be zero. If not, which in practice is the likely case, there will be errors. Current practice seems to minimize these errors by non-linear least squares, although other criteria may be employed.

<sup>2</sup>More specifically, the Jacobian of the transformation is given by differentiating (2).

can be shown (e.g. Chapter 7 in Duffie (2001)) that the following PDE must be satisfied in order to rule out arbitrage opportunities:

$$F_t + \tilde{\mu}(t, X_t) F_x + \frac{1}{2} \text{trace}(\sigma(t, X_t) \sigma(t, X_t)^\top F_{xx}) - r_t F = 0, \quad (7)$$

where  $F(t, X_t)$  is the price of an arbitrary claim,  $\tilde{\mu}(t, X_t)$  is the drift under the risk-neutral measure  $Q$ , and subscripts denote partial derivatives with respect to the argument.

Using the representation theorem of Feynman-Kac, it is possible to derive (1). For reasons of tractability, a particularly useful way of specifying (6) is to assume that zero-coupon bond prices should be exponential-affine functions of the state variables, that is

$$P(t, T) = e^{A(t, T) + B(t, T) \cdot X_t}. \quad (8)$$

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. 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 a different measure, given the following transformation:

$$dW_t^Q = dW_t - g_t dt, \quad (9)$$

where  $g_t$  is the kernel in the transformation. By substituting this into (6), we obtain the dynamics of the state variables under  $Q$ ,

$$dX_t = [\mu(t, X_t) + \sigma(t, X_t) g_t] dt + \sigma(t, X_t) dW_t^Q.$$

By letting  $\Lambda_t = -g_t$  denote the so-called "market price of risk", this can be shown to correspond exactly to the Girsanov transformation necessary to apply the Feynman-Kac representation theorem on the partial differential equation to obtain the solution expressing the term structure of interest rates. There is therefore a one-to-one relationship between the market price of risk and the Girsanov kernel.

Restricting our attention to the affine class of models, by applying Ito's formula on (8) and using the term structure equation (7), it is possible to show that the  $Q$ -dynamics for the price of a zero-coupon  $T$ -bond is given by

$$dP_t(t, T) = r_t P_t(t, T) dt + B(t, T) \sigma(t, X_t) P(t, T) dW_t^Q. \quad (10)$$

The compensation for risk can be inferred from the instantaneous expected rate of return on a zero-coupon bond under  $P$ , which can be obtained by substituting (9) into the above equation and using that  $\Lambda_t = -g_t$ , resulting in

$$dP_t(t, T) = (r_t + \Lambda_t B(t, T) \sigma(t, X_t)) P_t(t, T) dt + B(t, T) \sigma(t, X_t) P(t, T) dW_t. \quad (11)$$

Hence, under the assumption of no arbitrage, the required local rate of return is equal to the short rate plus a factor which is proportional to the local volatility of the bond,

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

where  $\sigma_t(P) \equiv B(t, T) \sigma(t, X_t) P(t, T)$  denotes the volatility of the bond. From this relation, it is easy to see intuitively why  $\Lambda_t$  is commonly referred to as the market price of risk. Rearranging, we see that  $\Lambda$  is the required rate of return in excess of the risk-free short rate, divided by the volatility,

$$\Lambda_t = \frac{\mu_t(P) - r_t}{\sigma_t(P)}. \quad (12)$$

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  $\mu$ ,  $\sigma$  and  $\Lambda$  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.

We conclude the discussion of measures by noting that yet another probability measure is of some interest in this context. The forward- $T$  measure (or forward neutral measure) is by construction such that it creates independence between the history of variables up to time  $T$  and the variables at  $T$ . This is very useful if we are interested in pricing assets where the payoff is potentially correlated with the discount factor, since the two can be separated under this measure. Another interesting feature is that the unbiased expectations hypotheses actually holds under the forward- $T$  measure. That is, the expected future short rate under the  $T$ -measure, is equal to the instantaneous forward rate for the relevant horizon (e.g. Chapter 19 in Björk (1998)). Thus, the difference between  $T$  and  $P$  captures the error made by interpreting forward rates as future short rates. The relationship between the Wiener processes under the  $Q$  measure and under the forward- $T$  measure is given by

$$dW^Q = B(t, T) \sigma(t, X_t) dt + dW^T.$$

### 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.

#### 3.1 The CIR interest rate process

In contrast to models based on no-arbitrage arguments, the CIR model is developed within an intertemporal general equilibrium setting. This allows a consistent characterization of the functional form of the risk premium. The assumptions underlying the CIR economy imply that 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,

$$dr_t = \kappa(\theta - r_t) dt + \sigma\sqrt{r_t}dW_t, \quad (13)$$

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  $\theta$  at an adjustment rate determined by  $\kappa$ , and that the variance of the short rate is proportional to the interest rate level.

#### 3.2 Risk premia in the CIR model

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  $\lambda r$  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  $\lambda$  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 following condition for bonds of any maturity:

$$\Lambda = \frac{\mu(P) - r}{\sigma(P)}.$$

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. Specifically, we know that according to the model the local volatility of a zero-coupon bond is (see Appendix C)

$$\sigma(P) = \frac{P_r \sigma(r)}{P} = \frac{P_r \sigma \sqrt{r}}{P}, \quad (14)$$

and that the market price of risk therefore is

$$\Lambda_{CIR} = \frac{\lambda \sqrt{r}}{\sigma}. \quad (15)$$

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

$$dr_t = \tilde{\kappa} (\tilde{\theta} - r_t) dt + \sigma \sqrt{r_t} d\tilde{W}_t, \quad (16)$$

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

$$\tilde{\kappa} = \kappa + \lambda, \quad (17)$$

$$\tilde{\theta} = \frac{\kappa \theta}{(\kappa + \lambda)}, \quad (18)$$

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  $\lambda$  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<sup>3</sup>,

$$\begin{aligned} \text{Var}[r(s) | r(t)] &= r(t) \left( \frac{\sigma^2}{\kappa} \right) (\exp[-\kappa(s-t)] - \exp[-2\kappa(s-t)]) \\ &\quad + \theta \left( \frac{\sigma^2}{2\kappa} \right) (1 - \exp[-\kappa(s-t)])^2. \end{aligned} \quad (19)$$

The impact on the conditional variance of a change in the risk parameter can be divided into two components. First, a larger (negative)  $\lambda$  leads to a slower speed of mean-reversion under  $Q$ , (17), 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)  $\lambda$  increases the risk neutral mean-reversion level, (18), thereby raising the variance (since it is increasing in the level of the short rate).

### 3.3 Interest rate and bond yield densities

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 maturities longer than instantaneous simply are affine functions of the short rate,

$$y(\tau) = a(\tau) + b(\tau)r(t), \quad (20)$$

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 ( $\tau$ ) 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 1 to calculate densities under  $P$  and under  $Q$ . The values in Table 1 were obtained by estimating the model with the methods used by Ait-Sahalia (1996) on German weekly term structure data between January 1996 and March 2002.

---

<sup>3</sup>See Cox et al. (1985), p. 392.

**Table 1:** GMM estimates of CIR parameters under the objective probability measure  $P$ , and fitted value of  $\lambda$

Discretized CIR model:  

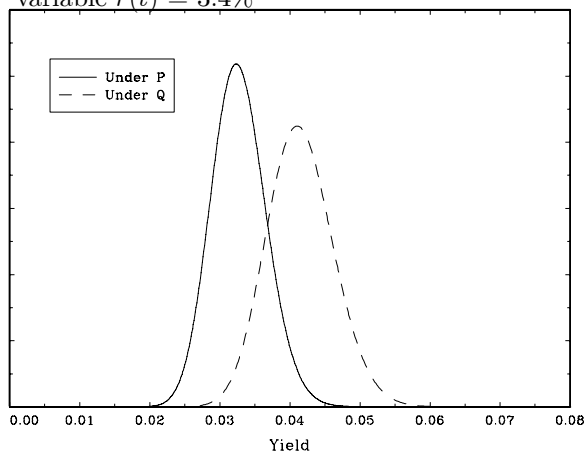
$$\Delta r_t = \kappa(\theta - r_t) \Delta t + \sigma \sqrt{r_t} \Delta W_t$$

Parameter	Estimate	st.err.
$\kappa$	0.523	0.257
$\theta$	0.031	0.005
$\sigma$	0.027	0.003
$\lambda$	-0.295	-

Standard errors are asymptotic, based on the optimal weighting matrix of Hansen (1982).

Chart 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%.<sup>4</sup> From the chart, 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  $\lambda$ , 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.

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



<sup>4</sup>3.4% is the level of the shortest interest rate at the end of the data sample. The resulting densities could therefore be seen as the densities implied by the CIR model at the end of March 2002.

The CIR model and our set of estimated parameters seem to generate substantial differences between conditional instantaneous interest rate densities under the risk-neutral measure and the corresponding densities under the actual probability measure, as displayed in Chart 1. An interesting issue is how this result carries over to the impact on conditional densities of yields on bonds with maturities longer than instantaneous when changing the probability measure. Particularly interesting, when considering our original problem of examining option-implied densities, would be to investigate the distributions of three-month and ten-year yields, since the most traded options in the euro area are based on underlying instruments of these maturities.

Since yields are affine in the instantaneous interest rate, (20), the yield densities will retain the functional form of the short rate density in Chart 1. The specific shape of a density for a given maturity  $\tau$  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  $\lambda$  will, apart from a Jensen's inequality term, determine how big these differences turn out to be.

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

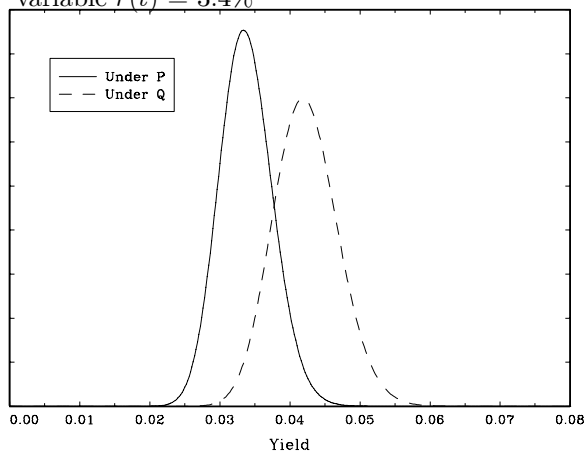
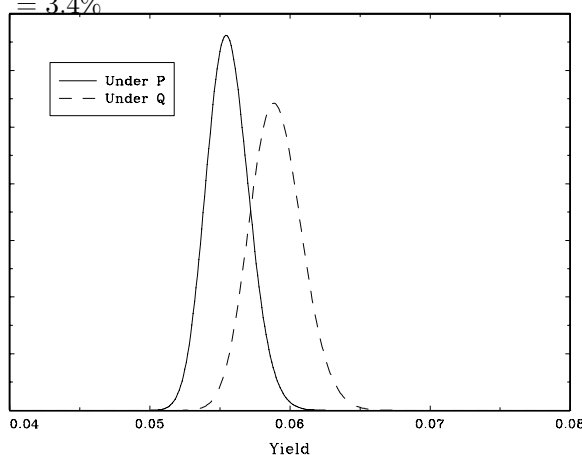


Chart 2 displays the one-year ahead distribution of the three-month interest rate conditional on an initial short rate of 3.4%. It is not surprising that the three-month densities resemble the instantaneous densities in Chart 1, given the relative similarity of the maturities. Greater differences are visible for the corresponding one-year ahead densities of ten-year bond yields in

Chart 3. First, the densities are centered around higher levels than the densities in Charts 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.

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



## 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.<sup>5</sup> 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  $X_t$ ,

$$r_t = \delta_0 + \delta'_x X_t, \quad (21)$$

and where  $X_t$  follows an affine diffusion,

$$\begin{aligned} dX_t &= \mathcal{K}(\Theta - X_t) dt + \Sigma \sqrt{S_t} dB_t \\ &\equiv \mu(X) dt + \sigma(X) dB_t, \end{aligned} \quad (22)$$

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

$$[S_t]_{ii} = \alpha_i + \beta'_i X_t, \quad (23)$$

where  $\alpha_i$  is a scalar and  $\beta_i$  is  $N \times 1$ . For future reference, we let  $\alpha$  denote the  $N$ -vector consisting of the individual  $\alpha_i$ s, and

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

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 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  $\mathcal{A}_m(N)$  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

---

<sup>5</sup>Dai and Singleton (2000) refer to an ATSM as admissible if it ensures positive factor variances, or, more precisely, 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 the minimal normalizations for econometric identification.

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  $X' = \left( X_{m \times 1}^{B'}, X_{(N-m) \times 1}^{D'} \right)$ . The parameter matrices and vectors are then partitioned accordingly, and the following normalizations are imposed:

$$\mathcal{K} = \begin{bmatrix} \mathcal{K}_{m \times m}^{BB} & \mathbf{0}_{m \times (N-m)} \\ \mathcal{K}_{(N-m) \times m}^{DB} & \mathcal{K}_{(N-m) \times (N-m)}^{DD} \end{bmatrix}, \quad (24)$$

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

$$\Theta = \begin{pmatrix} \Theta_{m \times 1}^B \\ \mathbf{0}_{(N-m) \times 1} \end{pmatrix}, \quad (25)$$

$$\Sigma = I, \quad (26)$$

$$\alpha = \begin{pmatrix} \mathbf{0}_{m \times 1} \\ \mathbf{1}_{(N-m) \times 1} \end{pmatrix}, \quad (27)$$

and

$$\mathcal{B} = \begin{bmatrix} I_{m \times m} & \mathcal{B}_{m \times (N-m)}^{BD} \\ \mathbf{0}_{(N-m) \times m} & \mathbf{0}_{(N-m) \times (N-m)} \end{bmatrix}. \quad (28)$$

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: $\mathcal{A}_0(3)$

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  $\mathcal{A}_0(3)$  model. In this case, the canonical representation sets  $\alpha$  equal to a  $3 \times 1$  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

$$\Lambda(t) = \sqrt{S_t} \lambda, \quad (29)$$

where  $\lambda$  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.<sup>6</sup> 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.<sup>7</sup> 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 (see Appendix D).<sup>8</sup>

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

$$\Lambda_t = \sqrt{S_t} \lambda + \sqrt{S_t^-} \psi X_t, \quad (30)$$

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

$$\sqrt{S_{t(ii)}^-} = \begin{cases} (\alpha_i + \beta_i' X_t)^{-1/2} & \text{if } \inf (\alpha_i + \beta_i' X_t) > 0 \\ 0 & \text{otherwise,} \end{cases} \quad (31)$$

while  $\psi$  is a  $3 \times 3$  parameter matrix. Given the canonical representation as specified in (24)-(28), it is easy to see why there is a trade-off between the flexibility of  $\sigma(X)$  and that of  $\Lambda$ . Clearly, in the  $\mathcal{A}_3(3)$  case, where by (27)  $\alpha = 0$  and by (28)  $\mathcal{B}$  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 (31), and  $\Lambda$  collapses to the completely affine specification (29). 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  $\psi X_t$  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  $\Lambda$  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

<sup>6</sup> ATSMs with this type of market price of risk are therefore sometimes referred to as "completely affine".

<sup>7</sup> Duarte (2003) 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.

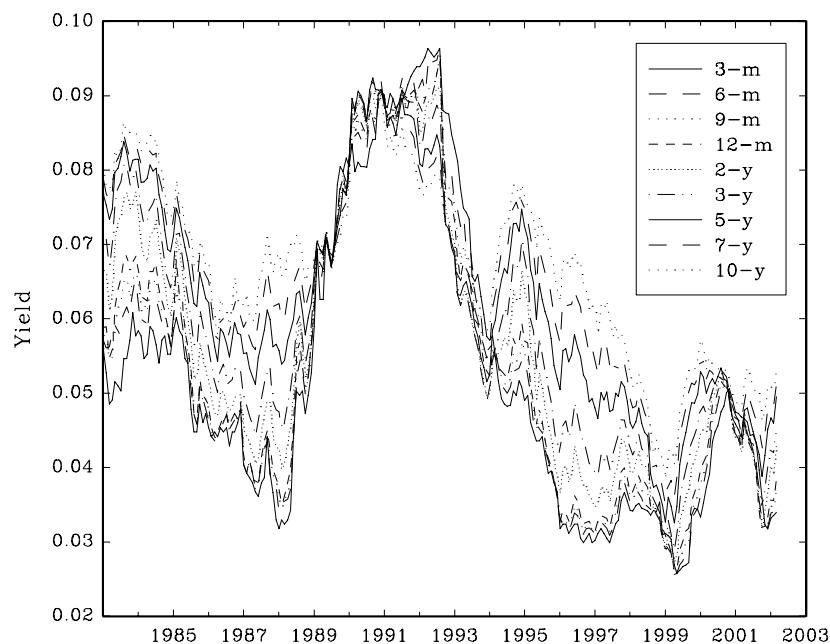
<sup>8</sup> The label "essentially affine" refers to the fact that while the factor processes remain affine under a change of measure, the variance of the state price deflator,  $\Lambda \Lambda'$ , is not affine. This is, however, inconsequential since this variance does not affect bond prices.



the most flexible specification of  $\Lambda$ , 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.

**Chart 4:** Yield data used in the estimations of multi-factor ATSMs



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 have therefore substituted 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) estimation. 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 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,  $Y_t = (y_t(\tau_1), \dots, y_t(\tau_M))$ , where

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

and  $\tau_i$  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,

$$Y_t = d(\phi) + Z(\phi) X_t + \varepsilon_t, \tag{32}$$

where the measurement error  $\varepsilon_t$  is assumed to be normally distributed,

$$\varepsilon_t \sim N(0, H(\phi)), \quad (33)$$

and  $\phi$  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 (32) with the theoretical expression for the yield on a zero-coupon bond with maturity  $\tau$  in the affine framework,

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

we see that

$$d(\phi) = -\frac{A(\tau)}{\tau} \quad (34)$$

and

$$Z(\phi) = -\frac{B(\tau)'}{\tau}. \quad (35)$$

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

$$\frac{dA(\tau)}{d\tau} = -\delta_0 - B'(\tau)\lambda + \frac{1}{2}B'(\tau)B(\tau), \quad (36)$$

$$\frac{dB(\tau)}{d\tau} = -\delta_y - \psi' B(\tau) - \mathcal{K}' B(\tau). \quad (37)$$

Here, we can note that since  $A(\tau)$  and  $B(\tau)$  are functions of, among other things, the  $\lambda$  and  $\psi$  parameter matrices in the market price of risk (30), 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,

$$X_t = \Phi(\phi) X_{t-h} + u_t, \quad (38)$$

where

$$\Phi(\phi) = \exp[-\mathcal{K}h], \quad (39)$$

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 (38) does not contain any constant in the  $\mathcal{A}_0(3)$  case since the normalization (25) in the canonical representation sets  $\Theta = 0$ . Because the model is Gaussian, the innovation in the transition equation is normally distributed,

$$u_t \sim N(0, V_t(\phi)), \quad (40)$$

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

$$V_t(\phi) = \int_{t-h}^t \exp[-\mathcal{K}(t-s)] \exp[-\mathcal{K}'(t-s)] ds. \quad (41)$$

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 E contains some further technical details on the implementation of the Kalman filter in this setting.<sup>9</sup>

**Table 2:** QML parameter estimates for the Essentially Affine  $\mathcal{A}_0(3)$  model

Parameter	Estimate	Standard error
$\delta_0$	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
$\kappa_{11}$	0.196	0.140
$\kappa_{21}$	-0.415	0.116
$\kappa_{22}$	1.363	0.204
$\kappa_{31}$	-0.314	0.334
$\kappa_{32}$	1.344	0.494
$\kappa_{33}$	0.151	0.080
$\lambda_1$	-0.405	0.219
$\lambda_2$	-0.052	0.032
$\lambda_3$	-1.469	0.267
$\psi_{11}$	-0.191	0.102
$\psi_{12}$	0.765	0.365
$\psi_{13}$	0.009	0.019
$\psi_{21}$	0.185	0.169
$\psi_{22}$	-0.151	0.186
$\psi_{23}$	-0.059	0.021
$\psi_{31}$	0.431	0.162
$\psi_{32}$	0.374	0.451
$\psi_{33}$	0.026	0.021

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.

Applying the estimation procedure outlined above, we obtain parameter estimates for the  $\mathcal{A}_0(3)$  model, as reported in Table 2. 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

<sup>9</sup>See also e.g. Lund (1997) or de Jong (2000) for further details on implementing the Kalman filter to estimate ATSMs.

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 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  $\tau$ -maturity zero-coupon yield at time  $t+h$  conditional on information at time  $t$  will be

$$\begin{aligned} E[y_{t+h}(\tau) | X_t] &= -\frac{A(\tau)}{\tau} - \frac{B(\tau)'}{\tau} E[X_{t+h} | X_t] \\ &= -\frac{A(\tau)}{\tau} - \frac{B(\tau)'}{\tau} \exp[-\mathcal{K}h] X_t, \end{aligned} \quad (42)$$

where the second equality follows from (38) and (39). Moreover, the corresponding conditional variance will be given by

$$\begin{aligned} Var[y_{t+h}(\tau) | X_t] &= \frac{B(\tau)'}{\tau} Var[X_{t+h} | X_t] \frac{B(\tau)}{\tau} \\ &= \frac{B(\tau)'}{\tau} \left( \int_t^{t+h} \exp[-\mathcal{K}(t-s)] \exp[-\mathcal{K}'(t-s)] ds \right) \frac{B(\tau)}{\tau}. \end{aligned} \quad (43)$$

Hence, the conditional objective ( $P$ ) distribution of  $y_{t+h}(\tau)$  will be Gaussian, with mean given by (42) and variance (43). The corresponding risk-neutral ( $Q$ ) distribution can be found by using the risk-neutral dynamics when evaluating the conditional expectation and variance of  $X$ :

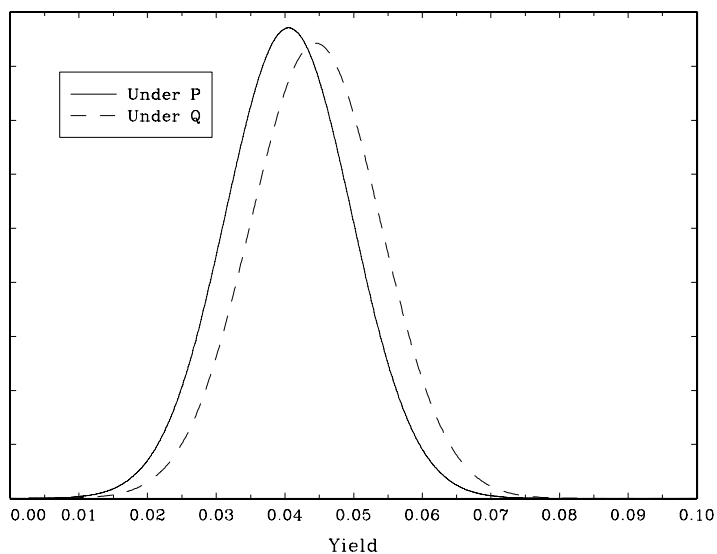
$$E^Q[X_{t+h} | X_t] = (I - \exp[-\mathcal{K}^Q h]) \Theta^Q + \exp[-\mathcal{K}^Q h] X_t, \quad (44)$$

$$Var^Q[X_{t+h} | X_t] = \int_t^{t+h} \exp[-\mathcal{K}^Q(t-s)] \exp[-\mathcal{K}^{Q'}(t-s)] ds, \quad (45)$$

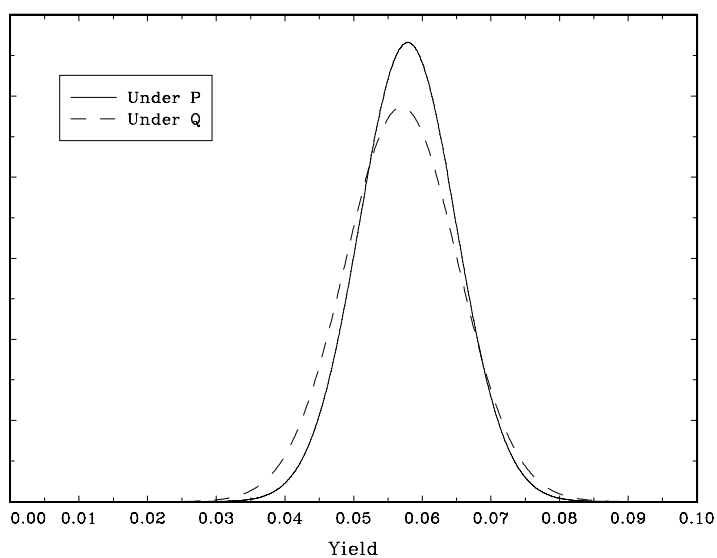
where  $\mathcal{K}^Q = (\mathcal{K} + \psi)$ , and  $\Theta^Q = -(\mathcal{K} + \psi)^{-1} \lambda$  follow from Girsanov's theorem.

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. Charts 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).

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



**Chart 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.



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.<sup>10</sup> 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%. Hence, by relying on e.g. option-implied densities, which deliver the risk-neutral probabilities, to infer the market's expectations about future outcomes, there is a considerable risk that incorrect conclusions will be drawn, at least if the assumed specification is the correct one.

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 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 (41) 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  $\mathcal{A}_1(3)$  model in the next section.

### 4.3 The $\mathcal{A}_1(3)$ essentially affine model

In the  $\mathcal{A}_1(3)$  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 (24)-(28). Notably, the first element of the vector of long-run means,  $\Theta$ , 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  $\mathcal{A}_1(3)$  specification on the conditional variance can easily be seen by comparing with the  $\mathcal{A}_0(3)$  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  $\mathcal{A}_1(3)$  model is given by

$$S_t = \begin{bmatrix} X_{1t} & 0 & 0 \\ 0 & 1 + \beta_{12}X_{1t} & 0 \\ 0 & 0 & 1 + \beta_{13}X_{1t} \end{bmatrix}. \quad (46)$$

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

<sup>10</sup>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.

$\inf (\alpha_1 + \beta_1' X_t)$  is no longer greater than zero, (31) requires that  $S_{t(11)}^- = 0$ , resulting in

$$\sqrt{S_t^-} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & (1 + \beta_{12} X_{1t})^{-1/2} & 0 \\ 0 & 0 & (1 + \beta_{13} X_{1t})^{-1/2} \end{bmatrix}. \quad (47)$$

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

As in the  $\mathcal{A}_0(3)$  case, we estimate the  $\mathcal{A}_1(3)$  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 (32) 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  $\Theta$  no longer has all elements equal to zero. Specifically, we get

$$X_t = c(\phi) + \Phi(\phi) X_{t-h} + u_t, \quad (48)$$

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

$$\begin{aligned} c(\phi) &= \int_{t-h}^t \exp[-\mathcal{K}(t-s)] \mathcal{K} \Theta ds \\ &= (I - \exp[-\mathcal{K}(h)]) \Theta, \end{aligned} \quad (49)$$

while  $\Phi(\phi)$  is defined as in (39). 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

$$V_t(\phi) = \int_{t-h}^t \exp[-\mathcal{K}(t-s)] S_t \exp[-\mathcal{K}'(t-s)] ds. \quad (50)$$



**Table 3:** QML parameter estimates for the Essentially Affine  $\mathcal{A}_1(3)$  model

Parameter	Estimate	Standard error
$\delta_0$	0.037	0.001
$\delta_1 \times 10^2$	0.102	0.006
$\delta_2 \times 10^2$	0.585	0.010
$\delta_3 \times 10^2$	0.139	0.043
$\theta_1$	7.351	0.054
$\kappa_{11}$	0.050	0.001
$\kappa_{21}$	-0.028	0.003
$\kappa_{22}$	0.284	0.019
$\kappa_{23}$	0.731	0.337
$\kappa_{31} \times 10^2$	-0.075	0.367
$\kappa_{32} \times 10^2$	-0.044	1.683
$\kappa_{33}$	1.252	0.085
$\beta_{12}$	0.221	0.006
$\beta_{13}$	1.046	0.061
$\lambda_1$	-0.018	0.003
$\lambda_2$	-0.278	0.006
$\lambda_3$	-0.006	0.002
$\psi_{21}$	0.042	0.018
$\psi_{22}$	0.005	0.013
$\psi_{23}$	0.381	0.382
$\psi_{31}$	0.212	0.028
$\psi_{32}$	-0.118	0.016
$\psi_{33}$	-0.201	0.129

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.

Equipped with parameter estimates obtained as outlined above, and as reported in Table 3, we can again go back and examine the differences between the risk-neutral and the objective yield distributions. 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  $x$  can be written

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

Now, 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))

$$\Pr_t^Q(y_T(\tau) \geq z) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Im} \left[ \psi^Q \left( -iv \frac{B(\tau)}{\tau}, X_t, t, T \right) \exp \left( -iv \left( z + \frac{A(\tau)}{\tau} \right) \right) \right]}{v} dv, \quad (51)$$

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

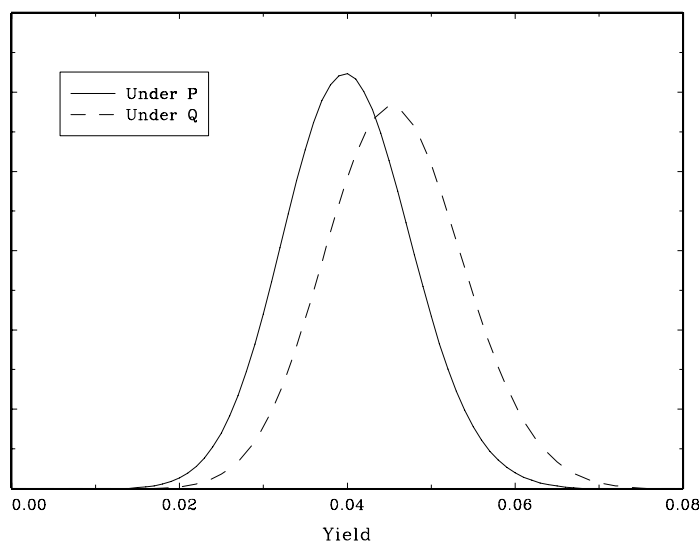
$$\psi^Q \left( -iv \frac{B(\tau)}{\tau}, X_t, t, T \right) = E_t^Q \left[ \exp \left( -iv \frac{B(\tau)}{\tau} X_T \right) \right] \quad (52)$$

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

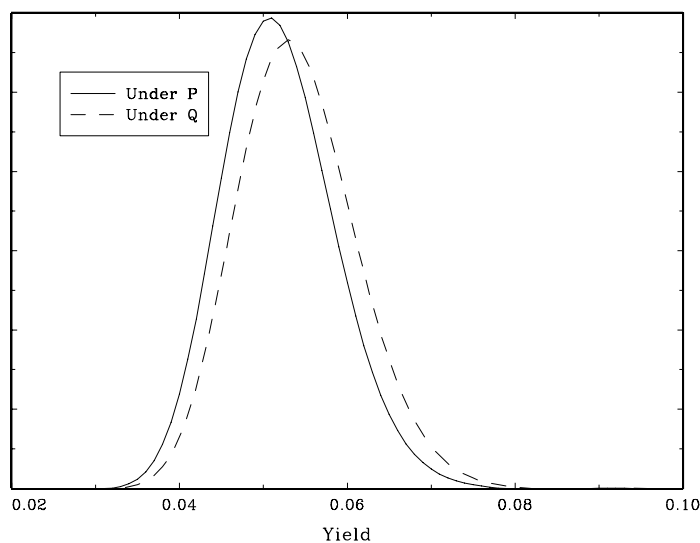
$$\psi^Q \left( -iv \frac{B(\tau)}{\tau}, X_t, t, T \right) = \exp(\alpha(t) + \beta(t) \cdot X_t), \quad (53)$$

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

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



**Chart 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.



Charts 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 Charts 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, 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 Chart 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 4 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.

**Table 4:** 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.

	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%

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.

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 Charts 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  $\mathcal{A}_1(3)$  model, the risk-neutral density is always centered to the right of the objective density (see Charts 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. Clearly, further analysis is needed to investigate this issue. In the meantime, the fact that two different  $\mathcal{A}(3)$ -specifications can produce such different results points to the need for extensive specification testing when choosing a specific model to be used in empirical applications.

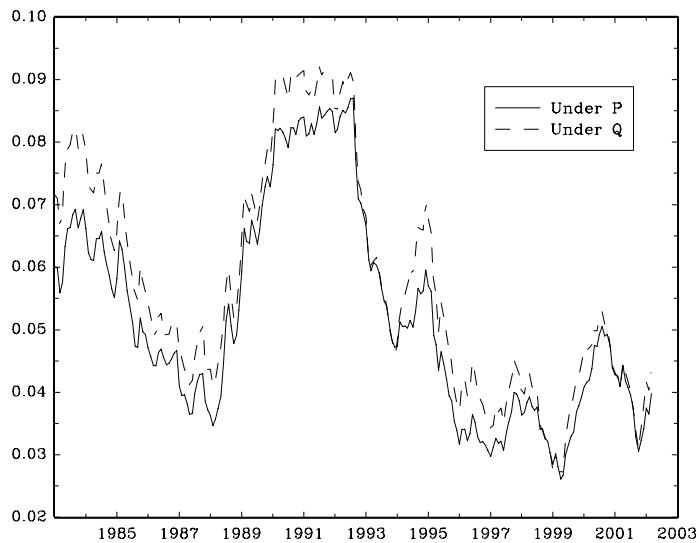
Second, in the example above which relates to end-March 2002, we see that the densities are more dispersed in the  $\mathcal{A}_0(3)$  case than in the  $\mathcal{A}_1(3)$  case. This is due to the fact that while the yield variance appeared to be relatively low in early 2002, the former model imposes 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 Charts 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.

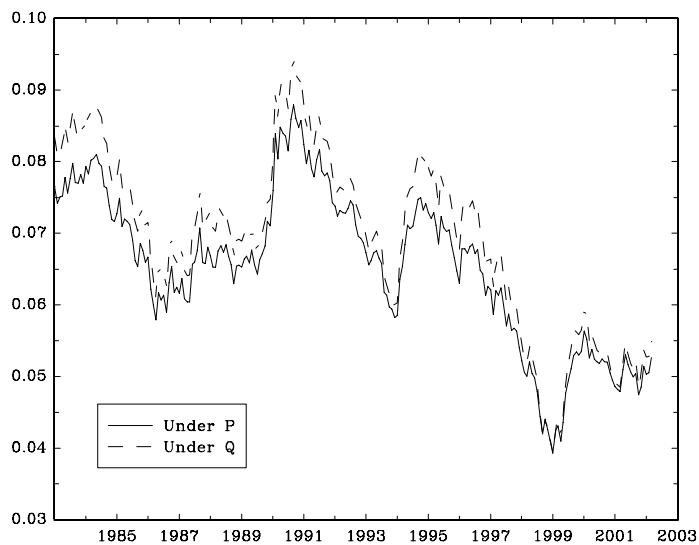
Third, in both cases, the dispersion of the 3-month PDFs is greater than for the 10-year counterparts. This is due to the typically downward-sloping term structure of yield volatilities observed in actual yield curve data. If a term structure model is successful in capturing this feature, one would expect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

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 Chart 9 and 10 with the corresponding 6-month and 3-month ahead PDFs displayed in Charts 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.

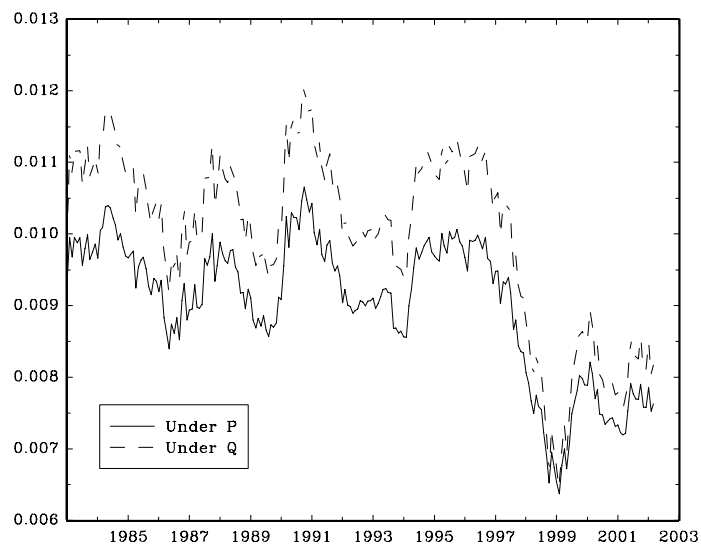
**Chart 9:** The mean of the one-year ahead conditional distribution of the 3-month interest rate, as implied by the  $\mathcal{A}_1(3)$ -model, January 1983 - March 2002.



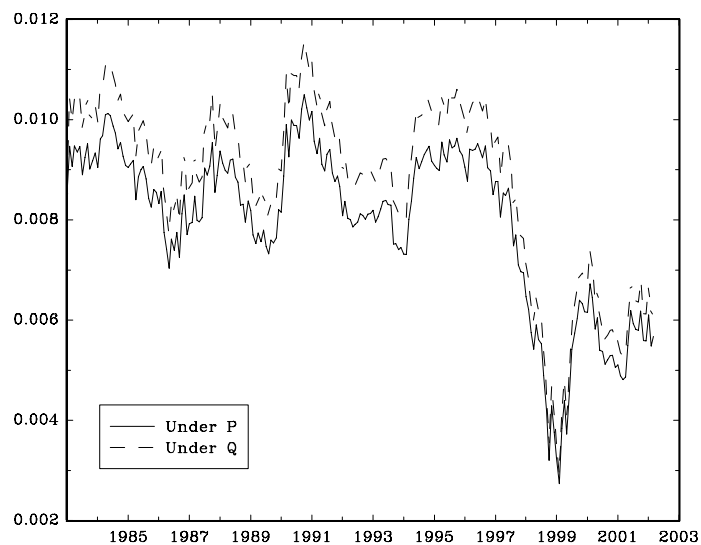
**Chart 10:** The mean of the one-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the  $\mathcal{A}_1(3)$ -model, January 1983 - March 2002.



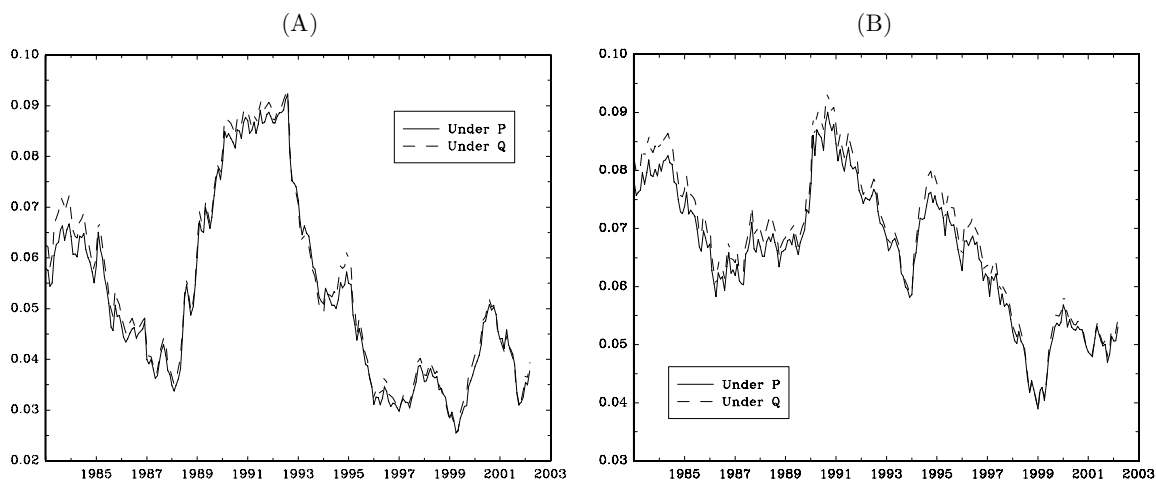
**Chart 11:** The square-root of the variance of the one-year ahead conditional distribution of the 3-month interest rate, as implied by the  $\mathcal{A}_1(3)$ -model, January 1983 - March 2002.



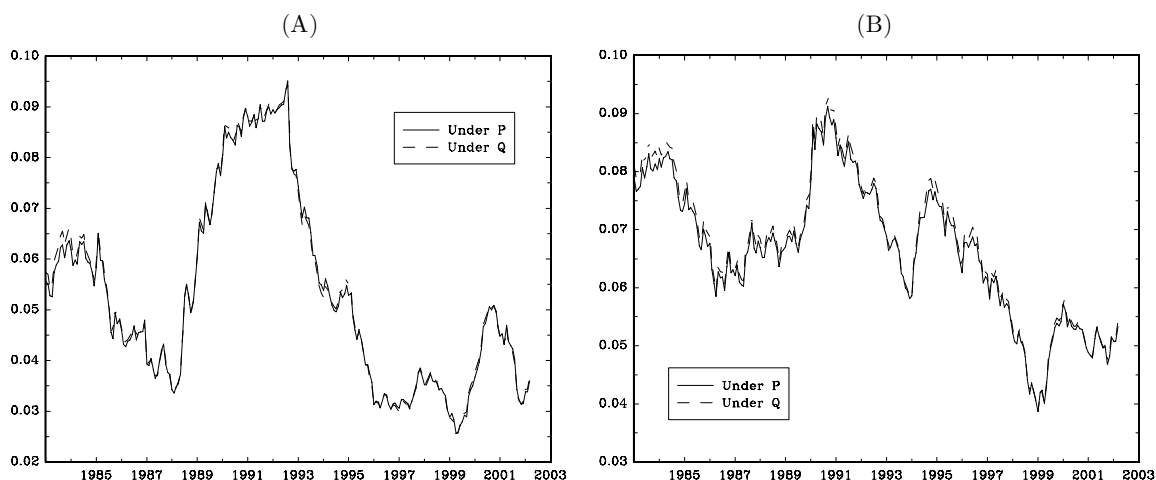
**Chart 12:** The square-root of the variance of the one-year ahead conditional distribution of the 10-year zero-coupon bond yield, as implied by the  $\mathcal{A}_1(3)$ -model, January 1983 - March 2002.



**Chart 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.



**Chart 13:** 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.



## 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.

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. For example, the maximum average yield error among the 9 yield series used in the estimation was 3.3 basis points for the  $\mathcal{A}_0(3)$  model and 2.4 basis points for the  $\mathcal{A}_1(3)$  model. Moreover, for both models, the root-mean-squared-errors ranged between 22 and 30 basis points across the 9 maturities. 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.

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

$$\begin{aligned} x_t &= \int_{-\infty}^{r_t} f_t(u) du \\ &= F_t(r_t). \end{aligned} \tag{54}$$

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,

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

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 (54) into a normal distribution (under the null). More precisely, if  $x_t$  is IID  $U(0, 1)$ , then the

transformed series

$$\begin{aligned} z_t &= \Phi^{-1} \left[ \int_{-\infty}^{r_t} f_t(u) du \right] \\ &= \Phi^{-1}(x_t) \end{aligned} \tag{55}$$

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.<sup>11</sup>

Table 5 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 reject in half of the cases, for both maturities. The critical element in these rejections seems to be that  $\mu_z$  differs significantly from zero, meaning that the conditional means of the estimated densities differ systematically from the realizations.

---

<sup>11</sup>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.

**Table 5:** Berkowitz (2001) density forecast evaluation tests

PANEL A: Evaluation of $\mathcal{A}_0(3)$ densities				
3-month forecast horizon				
	LR <sub>1</sub>	LR <sub>2</sub>	LR <sub>3</sub>	LR <sub>4</sub>
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 <sub>1</sub>	LR <sub>2</sub>	LR <sub>3</sub>	LR <sub>4</sub>
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 <sub>1</sub>	LR <sub>2</sub>	LR <sub>3</sub>	LR <sub>4</sub>
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 <sub>1</sub>	LR <sub>2</sub>	LR <sub>3</sub>	LR <sub>4</sub>
3m. rate / $P$ -measure	1.80	25.49*	26.65*	6.57
10y. rate / $P$ -measure	3.82	1.49	4.91	6.97

The test statistics are chi-squared distributed 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  $\rho$ .

Next, we turn to Panel B in Table 5, 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  $\mu_z$  from zero. This indicates 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.<sup>12</sup>

## 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 risk neutral densities, such as option-implied densities, as representing the true expectations of market participants. 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.

---

<sup>12</sup>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 true expectations.

## A Derivation of the ODEs for transform analysis

By definition, the characteristic function is

$$\phi_t^x(u, X_t, t, T) \equiv E^x \left( \exp \left( - \int_t^T R(X_s) ds \right) e^{iu'Y_T} | X_t \right),$$

where  $u \in R^n$ . In Duffie, Pan and Singleton (2000), this is seen in (2.3) p. 1351 where the above equation is defined without  $i$  but where  $u \in C^n$ . However, as will be clear from the derivation of the set of ODE:s,  $\text{Re}(u) = 0$ . In Singleton (2001), the special case of the above function is

$$\phi_t(u, Y_t, t, T) = E \left( e^{iu'Y_T} | Y_t \right). \quad (56)$$

The main thing that these papers prove is that if we are willing to impose an affine form for the evolution of the state variables (Chacko (1999) generalizes this to be affine in  $r$ , and exponential separable in the other variables), the characteristic function will take the form

$$\phi_t(u, X_t, t, T) = e^{\alpha(t) + \beta(t) \cdot x}$$

where  $\alpha(t)$ ,  $\beta(t)$  are complex valued functions that solve a set of differential equations. Examining (56), we see that if we fix  $u$  and  $T$ ,  $\phi_t$  will be a stochastic process where the dynamic properties will be determined by the behavior of  $Y_t$ . We see that when  $t = T$ , there is no random element and the definition reduces to  $e^{iu'Y_T}$ . In order for our candidate function to be correct, it must satisfy two properties. First, by the law of iterated expectations, the expected change must be equal to zero. Second, to comply with the deterministic terminal condition, we must have

$$e^{\alpha(t) + \beta(t) \cdot Y_T} = e^{iu'Y_T},$$

and therefore

$$\begin{aligned} \alpha(T) &= 0 \\ \beta(T) &= iu. \end{aligned}$$

To explicitly show the dependence of  $\alpha(t)$ ,  $\beta(t)$  on  $u$ , we write

$$\alpha_t(u).$$

By applying Ito's formula to the candidate function we get

$$\frac{\partial \phi_t}{\partial t} dt + \frac{\partial \phi_t}{\partial Y} dY_t + \frac{1}{2} \frac{\partial^2 \phi_t}{\partial Y^2} dY^2$$

By assumption,

$$dY_t = \mu dt + \sigma dW_t,$$

where  $\mu$  and  $\sigma^2$  are affine functions of the state variables. Further (illustrating for simplicity the scalar case),

$$\begin{aligned}\frac{\partial \phi_t}{\partial t} &= [\dot{\alpha}_t + \dot{\beta}_t Y_t] \phi_t \\ \frac{\partial \phi_t}{\partial Y} &= \beta_t \phi_t \\ \frac{\partial^2 \phi_t}{\partial Y^2} &= \beta_t^2 \phi_t.\end{aligned}$$

Thus,

$$\begin{aligned}(\dot{\alpha}_t + \dot{\beta}_t Y_t) \phi_t dt + \beta_t \phi_t dY_t + \frac{1}{2} \beta_t^2 \phi_t dY_t^2 \\ (\dot{\alpha}_t + \dot{\beta}_t Y_t) \phi_t dt + \beta_t \phi_t [\mu dt + \sigma dW_t] + \frac{1}{2} \beta_t^2 \phi_t \sigma^2 dt \\ \left[ \dot{\alpha}_t + \dot{\beta}_t Y_t + \beta_t \mu + \frac{1}{2} \beta_t^2 \sigma^2 \right] \phi_t dt + \beta_t \sigma \phi_t dW_t\end{aligned}$$

By inserting the linear structure of  $\mu$  and  $\sigma$ , we get

$$\begin{aligned}\left[ \dot{\alpha}_t + \dot{\beta}_t Y_t + \beta_t (k_0 + k_1 Y_t) + \frac{1}{2} \beta_t^2 (h_0 + h_1 Y_t) \right] \phi_t dt + \beta_t \sigma \phi_t dW_t \\ \left[ \left( \dot{\alpha}_t + \beta_t k_0 + \frac{1}{2} h_0 \beta_t^2 \right) + \left( \dot{\beta}_t + \beta_t k_1 + \frac{1}{2} \beta_t^2 h_1 \right) Y_t \right] \phi_t dt + \beta_t \sigma \phi_t dW_t.\end{aligned}$$

By integrating this we get

$$\int_t^s \left[ \left( \dot{\alpha}_v + \beta_v k_0 + \frac{1}{2} h_0 \beta_v^2 \right) + \left( \dot{\beta}_v + \beta_v k_1 + \frac{1}{2} \beta_v^2 h_1 \right) Y_v \right] \phi_v dv + \int_t^s \beta_v \sigma \phi_v dW_v.$$

Finally, by using that  $E \int_t^s \frac{\partial \phi_v}{\partial (t, Y)} dv = 0$  by the law of iterated expectations, as well as the fact that  $dW$  is a Wiener process, we get

$$E \left[ \int_t^s \left[ \left( \dot{\alpha}_v + \beta_v k_0 + \frac{1}{2} h_0 \beta_v^2 \right) + \left( \dot{\beta}_v + \beta_v k_1 + \frac{1}{2} \beta_v^2 h_1 \right) Y_v \right] \phi_v dv | Y_t \right] = 0$$

In order for this to hold, we must have

$$\begin{aligned}\dot{\alpha}_t + \beta_t k_0 + \frac{1}{2} h_0 \beta_t^2 &= 0 \\ \dot{\beta}_t + \beta_t k_1 + \frac{1}{2} \beta_t^2 h_1 &= 0.\end{aligned}$$

## B The link between $r$ - and $P$ -distributions in a one-factor setting

In the scalar affine model, the link between prices and the states can be written

$$P(s, T; r_s) = e^{(A(s, T) - B(s, T)r)}.$$

Thus, standing at  $t$ , looking forward towards  $s$ , at a bond expiring at  $T$ , this is a random variable. However, given that we can find the distribution of  $x_s$  conditional on information at  $t$ , we can obtain the distribution of  $P(s, T; x_s)$  through the relation

$$\begin{aligned} f_P(P) &= f_{r_s}(P^{-1}(r_s)) \left| \frac{dr}{dP} \right| \\ P^{-1}(r_s) &= \frac{A(s, T) - \ln P}{B(s, T)} \\ \frac{dr}{dP} &= \frac{-1}{B(s, T)P}, \end{aligned}$$

thus,

$$f_P(P) = f_{r_s} \left( \frac{A(s, T) - \ln P}{B(s, T)} \right) \left| \frac{-1}{B(s, T)P} \right|.$$

When the bond price is expressed as

$$P(s, T; r_s) = A(s, T) e^{-B(s, T)r},$$

as in CIR (1985), we get

$$\begin{aligned} P^{-1}(r_s) &= -\frac{\ln \left( \frac{P}{A(s, T)} \right)}{B(s, T)}, \\ \frac{dr}{dP} &= -\frac{1}{B(s, T)P}, \end{aligned}$$

and

$$f_P(P) = f_{r_s} \left( -\frac{\ln \left( \frac{P}{A(s, T)} \right)}{B(s, T)} \right) \left| \frac{-1}{B(s, T)P} \right|.$$

To obtain the distribution of the zero-coupon yield associated with  $f_P(P)$  we need an analogous transformation. Letting  $f_{y_\tau}(y_\tau)$  denote the distribution of the  $\tau$ -horizon yield to maturity, we get

$$f_{y_\tau}(y_\tau) = f_{P_\tau}(y^{-1}(P_\tau)) \left| \frac{dP}{dy} \right|,$$

where

$$y^{-1}(P_\tau) = \exp(-\tau y_\tau),$$

and

$$\frac{dP}{dy} = -\tau \exp(-\tau y_\tau),$$

so that

$$f_{y_\tau}(y_\tau) = f_{P_\tau}(\exp(-\tau y_\tau)) |-\tau \exp(-\tau y_\tau)|.$$

We can rewrite this as a function of the distribution of the short rate  $r$  in the following way:

$$\begin{aligned} f_{y_\tau}(y_\tau) &= f_{r_s} \left( -\frac{\ln \left( \frac{\exp(-\tau y_\tau)}{A(s, T)} \right)}{B(s, T)} \right) \left| \frac{-1}{B(s, T) \exp(-\tau y_\tau)} \right| |-\tau \exp(-\tau y_\tau)| \\ &= f_{r_s} \left( \frac{\tau y_\tau + \ln A(s, T)}{B(s, T)} \right) \left| \frac{\tau}{B(s, T)} \right|. \end{aligned}$$

## C The bond price process in CIR

In the CIR model, the process for the only factor, the short-term interest rate, is given by the following affine diffusion:

$$\begin{aligned} dr &= \mu(r) dt + \sigma(r) dW \\ &= \kappa(\theta - r) dt + \sigma\sqrt{r}dW, \end{aligned}$$

and the price of a zero coupon bond price of maturity  $\tau$  is a function of the short rate and maturity-specific functions  $A(\tau)$  and  $B(\tau)$  according to

$$P(\tau) = A(\tau) \exp[-B(\tau)r].$$

The process for the bond price obeys

$$dP = \mu(P) dt + \sigma(P) dW,$$

where the functions  $\mu(P)$  and  $\sigma(P)$  can be determined by applying Ito's lemma:

$$\begin{aligned} \mu(P) &= -P_\tau + P_r\mu(r) + \frac{1}{2}\sigma^2(r)P_{rr}, \\ \sigma(P) &= P_r\sigma(r), \end{aligned}$$

where subscripts denote partial derivatives. Solving for the derivatives (notationally suppressing  $\tau$  as an argument) gives

$$\begin{aligned} P_\tau &= A_\tau \exp[-Br] - AB_\tau r \exp[-Br] \\ &= \left(\frac{A_\tau}{A} - B_\tau r\right) P, \\ P_r &= -AB \exp[-Br] \\ &= -BP, \\ P_{rr} &= B^2P, \end{aligned}$$

so that

$$\mu(P) = P \left[ \left( B_\tau r - \frac{A_\tau}{A} \right) - B\kappa(\theta - r) + \frac{1}{2}B^2\sigma^2 r \right],$$

and

$$\sigma(P) = -BP\sigma\sqrt{r}.$$

We can evaluate the  $\tau$ -derivatives of the functions  $A$  and  $B$  using the closed-form expressions for these functions:

$$\begin{aligned} A(\tau) &= \left[ \frac{\phi_1 \exp[\phi_2\tau]}{\phi_2 (\exp[\phi_1\tau] - 1) + \phi_1} \right]^{\phi_3}, \\ B(\tau) &= \frac{\exp[\phi_1\tau] - 1}{\phi_2 (\exp[\phi_1\tau] - 1) + \phi_1}, \end{aligned}$$



where

$$\begin{aligned}\phi_1 &= \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}, \\ \phi_2 &= \frac{\kappa + \lambda + \phi_1}{2}, \\ \phi_3 &= \frac{2\kappa\theta}{\sigma^2},\end{aligned}$$

obtaining :

$$A_\tau = \frac{A\phi_2\phi_3 [\phi_1 (1 - \exp[\phi_1\tau]) + \phi_2 (\exp[\phi_1\tau] - 1)]}{\phi_2 (\exp[\phi_1\tau] - 1) + \phi_1},$$

and

$$B_\tau = \frac{\phi_1^2 \exp[\phi_1\tau]}{[\phi_2 (\exp[\phi_1\tau] - 1) + \phi_1]^2}.$$

## D Risk-neutral dynamics of $Y_t$ in the $\mathcal{A}_0(3)$ case

Given an assumed market price of risk vector, we can express the dynamics of  $Y_t$  under the risk-neutral probability measure  $Q$  as

$$dY_t = \mathcal{K}(\Theta - Y_t) dt - \Sigma\sqrt{S_t}\Lambda_t dt + \Sigma\sqrt{S_t}dB_t^Q.$$

For the essentially affine  $\mathcal{A}_0(3)$  model, the  $Q$ -dynamics are given by

$$\begin{aligned}dY_t &= [\mathcal{K}(\Theta - Y_t) - \Lambda_t] dt + \Sigma\sqrt{S_t}dB_t^Q \\ &= [\mathcal{K}(\Theta - Y_t) - \lambda - \psi Y_t] dt + \Sigma\sqrt{S_t}dB_t^Q \\ &= [(\mathcal{K}\Theta - \lambda) - (\mathcal{K} + \psi)Y_t] dt + \Sigma\sqrt{S_t}dB_t^Q \\ &= [-\lambda - (\mathcal{K} + \psi)Y_t] dt + \Sigma\sqrt{S_t}dB_t^Q,\end{aligned}$$

since  $\mathcal{K}\Theta = 0$ . Defining  $\mathcal{K}^Q\Theta^Q \equiv -\lambda$  and  $\mathcal{K}^Q \equiv (\mathcal{K} + \psi)$ , we can write

$$dY_t = (\mathcal{K}^Q\Theta^Q - \mathcal{K}^Q Y_t) dt + \Sigma\sqrt{S_t}dB_t^Q,$$

and we see that the risk neutral dynamics of  $Y$  remain affine.

## E The Kalman filter in the $\mathcal{A}_0(3)$ case

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  $\phi$ )

$$\begin{aligned}\hat{Y}_{t|t-h} &= E[Y_t | R_{t-h}] \\ &= \Phi\hat{Y}_{t-h},\end{aligned}\tag{57}$$

and an associated MSE matrix,

$$\begin{aligned}\hat{\Omega}_{t|t-h} &= E \left[ \left( Y_t - \hat{Y}_{t|t-h} \right) \left( Y_t - \hat{Y}_{t|t-h} \right)' \mid R_{t-h} \right] \\ &= \Phi \hat{\Omega}_{t-h} \Phi' + V_t.\end{aligned}\tag{58}$$

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

$$\begin{aligned}\hat{Y}_t &= E [Y_t \mid R_t] \\ &= \hat{Y}_{t|t-h} + \hat{\Omega}_{t|t-h} Z' F_t^{-1} v_t,\end{aligned}\tag{59}$$

where  $v_t$  is the vector of prediction errors,

$$v_t = R_t - \left( d + Z \hat{Y}_{t|t-h} \right),\tag{60}$$

and  $F_t$  is the covariance matrix of the prediction errors,

$$\begin{aligned}F_t &= \text{cov} (R_t \mid R_{t-h}) = E [v_t v_t' \mid R_{t-h}] \\ &= Z \hat{\Omega}_{t|t-h} Z' + H.\end{aligned}\tag{61}$$

The update of the MSE matrix  $\Omega$  is given by

$$\begin{aligned}\hat{\Omega}_t &= E \left[ \left( Y_t - \hat{Y}_t \right) \left( Y_t - \hat{Y}_t \right)' \mid R_t \right] \\ &= \hat{\Omega}_{t|t-h} - \hat{\Omega}_{t|t-h} Z' F_t^{-1} Z \hat{\Omega}_{t|t-h} \\ &= \left( I - \hat{\Omega}_{t|t-h} Z' F_t^{-1} Z \right) \hat{\Omega}_{t|t-h}.\end{aligned}\tag{62}$$

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:

$$\begin{aligned}\ln L &= \sum_{t=1}^T \ln L_t, \\ \ln L_t &= -\frac{1}{2} \ln (2\pi) - \frac{1}{2} \ln |F_t| - \frac{1}{2} v_t' F_t^{-1} v_t.\end{aligned}\tag{63}$$

## References

- [1] Ait-Sahalia, Y. (1996). "Nonparametric Pricing of Interest Rate Derivative Securities", *Econometrica*, 64, 527-560.
- [2] Ait-Sahalia, Y., Wang, Y., Yared, F. (2001). "Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?", *Journal of Econometrics*, 102, 67-110.
- [3] Björk, T. (1998). *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford, U.K.
- [4] Black, F., Scholes, M. (1973). "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 659-683.
- [5] Bliss, R.R., Panigirtzoglou, N., (2001). "Recovering Risk Aversion from Options", forthcoming, *Journal of Finance*.
- [6] Chacko, G., (1999). "Continuous-Time Estimation of Exponential Separable Term Structure Models: A General Approach", manuscript, Harvard University.
- [7] Cox, J., Ingersoll, J., Ross, S., (1985). "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-408.
- [8] Dai, Q., Singleton, K. (2000). "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, 55, 1943-1978.
- [9] Dai, Q., Singleton, K. (2002). "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure", *Journal of Financial Economics*, 63, 415-441.
- [10] de Jong, F. (2000). "Time Series and Cross-Section Information in Affine Term-Structure Models", *Journal of Business & Economic Statistics*, 18, 300-314.
- [11] Duarte, J. (2003). "Evaluating an Alternative Risk Preference in Affine Term Structure Models", working paper, forthcoming, *Review of Financial Studies*.
- [12] Duffee, G. (2002). "Term Premia and Interest Rate Forecasts in Affine Models", *Journal of Finance*, 57, 405-443.
- [13] Duffee, G., Stanton, R. (2001). "Estimation of Dynamic Term Structure Models", Working paper, U.C. Berkeley.
- [14] Duffie, D. (2001). *Dynamic Asset Pricing Theory*. (3rd edition) Princeton University Press, Princeton, N.J.
- [15] Duffie, D., Kan, R. (1996). "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, 6, 379-406.
- [16] Duffie, D., Singleton, K., (1993). "Simulated Moments Estimation of Markov Models of Asset Prices", *Econometrica*, 61, 929-952.
- [17] Duffie, D., Pan, J., Singleton, K., (2000). "Transform Analysis and Asset Pricing for Affine Jump-Diffusions", *Econometrica*, 68, 1343-1376.
- [18] Gallant, A.R., Tauchen, G.E. (1996). "Which Moments to Match?", *Econometric Theory*, 12, 657-681.
- [19] 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.
- [20] Lund, J., (1997). "Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates", Working paper, Aarhus School of Business.
- [21] Melick, W.R., 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.

- [22] Nelson, C.R., Siegel, A.F. (1987). "Parsimonious Modeling of Yield Curves", *Journal of Business*, 60, 473-489.
- [23] Singleton, K. (2001). "Estimation of Affine Asset Pricing Models using the Empirical Characteristic Function", *Journal of Econometrics*, 102, 111-141.
- [24] Söderlind, P., Svensson, L.E.O. (1997). "New Techniques to Extract Market Expectations from Financial Instruments", *Journal of Monetary Economics* 40 383-430.
- [25] Svensson, L.E.O. (1994). "Estimating and Interpreting Forward Interest Rates: Sweden 1992-94", IMF Working paper WP/94/114, Washington D.C.
- [26] Vasicek, O., (1977). "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, 177-188.
- [27] White, H., (1982). "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1-25.

## European Central Bank working paper series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (<http://www.ecb.int>).

- 202 "Aggregate loans to the euro area private sector" by A. Calza, M. Manrique and J. Sousa, January 2003.
- 203 "Myopic loss aversion, disappointment aversion and the equity premium puzzle" by D. Fielding and L. Stracca, January 2003.
- 204 "Asymmetric dynamics in the correlations of global equity and bond returns" by L. Cappiello, R.F. Engle and K. Sheppard, January 2003.
- 205 "Real exchange rate in an inter-temporal n-country-model with incomplete markets" by B. Mercereau, January 2003.
- 206 "Empirical estimates of reaction functions for the euro area" by D. Gerdesmeier and B. Roffia, January 2003.
- 207 "A comprehensive model on the euro overnight rate" by F. R. Würtz, January 2003.
- 208 "Do demographic changes affect risk premiums? Evidence from international data" by A. Ang and A. Maddaloni, January 2003.
- 209 "A framework for collateral risk control determination" by D. Cossin, Z. Huang, D. Aunon-Nerin and F. González, January 2003.
- 210 "Anticipated Ramsey reforms and the uniform taxation principle: the role of international financial markets" by S. Schmitt-Grohé and M. Uribe, January 2003.
- 211 "Self-control and savings" by P. Michel and J.P. Vidal, January 2003.
- 212 "Modelling the implied probability of stock market movements" by E. Glatzer and M. Scheicher, January 2003.
- 213 "Aggregation and euro area Phillips curves" by S. Fabiani and J. Morgan, February 2003.
- 214 "On the selection of forecasting models" by A. Inoue and L. Kilian, February 2003.
- 215 "Budget institutions and fiscal performance in Central and Eastern European countries" by H. Gleich, February 2003.
- 216 "The admission of accession countries to an enlarged monetary union: a tentative assessment" by M. Ca'Zorzi and R. A. De Santis, February 2003.
- 217 "The role of product market regulations in the process of structural change" by J. Messina, March 2003.

- 218 “The zero-interest-rate bound and the role of the exchange rate for monetary policy in Japan” by G. Coenen and V. Wieland, March 2003.
- 219 “Extra-euro area manufacturing import prices and exchange rate pass-through” by B. Anderton, March 2003.
- 220 “The allocation of competencies in an international union: a positive analysis” by M. Ruta, April 2003.
- 221 “Estimating risk premia in money market rates” by A. Durré, S. Evjen and R. Pilegaard, April 2003.
- 222 “Inflation dynamics and subjective expectations in the United States” by K. Adam and M. Padula, April 2003.
- 223 “Optimal monetary policy with imperfect common knowledge” by K. Adam, April 2003.
- 224 “The rise of the yen vis-à-vis the (“synthetic”) euro: is it supported by economic fundamentals?” by C. Osbat, R. Ruffer and B. Schnatz, April 2003.
- 225 “Productivity and the (“synthetic”) euro-dollar exchange rate” by C. Osbat, F. Visselaar and B. Schnatz, April 2003.
- 226 “The central banker as a risk manager: quantifying and forecasting inflation risks” by L. Kilian and S. Manganelli, April 2003.
- 227 “Monetary policy in a low pass-through environment” by T. Monacelli, April 2003.
- 228 “Monetary policy shocks – a nonfundamental look at the data” by M. Klaeffer, May 2003.
- 229 “How does the ECB target inflation?” by P. Surico, May 2003.
- 230 “The euro area financial system: structure, integration and policy initiatives” by P. Hartmann, A. Maddaloni and S. Manganelli, May 2003.
- 231 “Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero” by G. Coenen, A. Orphanides and V. Wieland, May 2003.
- 232 “Describing the Fed’s conduct with Taylor rules: is interest rate smoothing important?” by E. Castelnuovo, May 2003.
- 233 “The natural real rate of interest in the euro area” by N. Giammarioli and N. Valla, May 2003.
- 234 “Unemployment, hysteresis and transition” by M. León-Ledesma and P. McAdam, May 2003.
- 235 “Volatility of interest rates in the euro area: evidence from high frequency data” by N. Cassola and C. Morana, June 2003.

- 236 “Swiss monetary targeting 1974-1996: the role of internal policy analysis” by G. Rich, June 2003.
- 237 “Growth expectations, capital flows and international risk sharing” by O. Castrén, M. Miller and R. Stiegert, June 2003.
- 238 “The impact of monetary union on trade prices” by R. Anderton, R. E. Baldwin and D. Taglioni, June 2003.
- 239 “Temporary shocks and unavoidable transitions to a high-unemployment regime” by W. J. Denhaan, June 2003.
- 240 “Monetary policy transmission in the euro area: any changes after EMU?” by I. Angeloni and M. Ehrmann, July 2003.
- 241 Maintaining price stability under free-floating: a fearless way out of the corner?” by C. Detken and V. Gaspar, July 2003.
- 242 “Public sector efficiency: an international comparison” by A. Afonso, L. Schuknecht and V. Tanzi, July 2003.
- 243 “Pass-through of external shocks to euro area inflation” by E. Hahn, July 2003.
- 244 “How does the ECB allot liquidity in its weekly main refinancing operations? A look at the empirical evidence” by S. Ejerskov, C. Martin Moss and L. Stracca, July 2003.
- 245 “Money and payments: a modern perspective” by C. Holthausen and C. Monnet, July 2003.
- 246 “Public finances and long-term growth in Europe – evidence from a panel data analysis” by D. R. de Ávila Torrijos and R. Strauch, July 2003.
- 247 “Forecasting euro area inflation: does aggregating forecasts by HICP component improve forecast accuracy?” by K. Hubrich, August 2003.
- 248 “Exchange rates and fundamentals” by C. Engel and K. D. West, August 2003.
- 249 “Trade advantages and specialisation dynamics in acceding countries” by A. Zaghini, August 2003.
- 250 “Persistence, the transmission mechanism and robust monetary policy” by I. Angeloni, G. Coenen and F. Smets, August 2003.
- 251 “Consumption, habit persistence, imperfect information and the lifetime budget constraint” by A. Willman, August 2003.
- 252 “”Interpolation and backdating with a large information set” by E. Angelini, J. Henry and M. Marcellino, August 2003.
- 253 “Bond market inflation expectations and longer-term trends in broad monetary growth and inflation in industrial countries, 1880-2001” by W. G. Dewald, September 2003.

- 254 “Forecasting real GDP: what role for narrow money?” by C. Brand, H.-E. Reimers and F. Seitz, September 2003.
- 255 “Is the demand for euro area M3 stable?” by A. Bruggeman, P. Donati and A. Warne, September 2003.
- 256 “Information acquisition and decision making in committees: a survey” by K. Gerling, H. P. Grüner, A. Kiel and E. Schulte, September 2003.
- 257 “Macroeconomic modelling of monetary policy” by M. Klaeffling, September 2003.
- 258 “Interest rate reaction functions and the Taylor rule in the euro area” by P. Gerlach-Kristen, September 2003.
- 259 “Implicit tax co-ordination under repeated policy interactions” by M. Catenaro and J.-P. Vidal, September 2003.
- 260 “Aggregation-theoretic monetary aggregation over the euro area, when countries are heterogeneous” by W. A. Barnett, September 2003.
- 261 “Why has broad money demand been more stable in the euro area than in other economies? A literature review” by A. Calza and J. Sousa, September 2003.
- 262 “Indeterminacy of rational expectations equilibria in sequential financial markets” by P. Donati, September 2003.
- 263 “Measuring contagion with a Bayesian, time-varying coefficient model” by M. Ciccarelli and A. Rebucci, September 2003.
- 264 “A monthly monetary model with banking intermediation for the euro area” by A. Bruggeman and M. Donnay, September 2003.
- 265 “New Keynesian Phillips Curves: a reassessment using euro-area data” by P. McAdam and A. Willman, September 2003.
- 266 “Finance and growth in the EU: new evidence from the liberalisation and harmonisation of the banking industry” by D. Romero de Ávila, September 2003.
- 267 “Comparing economic dynamics in the EU and CEE accession countries” by R. Süppel, September 2003.
- 268 “The output composition puzzle: a difference in the monetary transmission mechanism in the euro area and the US” by I. Angeloni, A. K. Kashyap, B. Mojon and D. Terlizzese, September 2003.
- 269 “Zero lower bound: is it a problem with the euro area?” by G. Coenen, September 2003.
- 270 “Downward nominal wage rigidity and the long-run Phillips curve: simulation-based evidence for the euro area” by G. Coenen, September 2003.
- 271 “Indeterminacy and search theory” by N. Giammarioli, September 2003.



- 272 "Inflation targets and the liquidity trap" by M. Klaeffling and V. Lopez Perez, September 2003.
- 273 "Definition of price stability, range and point inflation targets: the anchoring of long-term inflation expectations" by E. Castelnuovo, S. Nicoletti-Altimari and D. Rodriguez-Palenzuela, September 2003.
- 274 "Interpreting implied risk neutral densities: the role of risk premia" by P. Hördahl and D. Vestin, September 2003.