Pastorello, Sergio

Working Paper

Multivariate Estimation of Exponential Affine Models of the Term Structure of Interest Rates

Quaderni - Working Paper DSE, No. 238

Provided in Cooperation with:
University of Bologna, Department of Economics

Suggested Citation: Pastorello, Sergio (1995) : Multivariate Estimation of Exponential Affine Models of the Term Structure of Interest Rates, Quaderni - Working Paper DSE, No. 238, Alma Mater Studiorum - Università di Bologna, Dipartimento di Scienze Economiche (DSE), Bologna, https://doi.org/10.6092/unibo/amsacta/5075

This Version is available at:
https://hdl.handle.net/10419/159081

Multivariate Estimation of Exponential Affine Models of the Term Structure of Interest Rates*

Sergio Pastorello†

February 29, 1996

Abstract

In this paper I consider the estimation of multi-factor exponential affine models of the term structure of interest rates. I start with a survey of the empirical work on the term structure in continuous time, showing that in most cases the implementation of the models has not fully exploited the theoretical restrictions. I also show that these works have almost always focused on "generalizations" of the theoretical model, based on the inclusion of measurement errors in bills and bonds prices. I then suggest two approaches to statistical inference: the first is based on the Kalman filter, while the second follows the indirect inference approach. I also briefly discuss the relative properties of the two estimators, and I conclude with a small Monte Carlo experiment for a one-factor Cox-Ingersoll-Ross model, whose results are rather encouraging.

*I wish to thank Carlo Bianchi, Flavio Cocco and Nizar Touzi for helpful comments and suggestions. I am responsible of any remaining error.

†Dipartimento di Scienze Economiche, Università di Bologna, Strada Maggiore 45, 40125 Bologna, ITALY. Ph: +39-51-6402622; Fax: +39-51-6402664; e-mail: pastorel@boph01.cineca.it .

1 Introduction

The term structure of interest rates has traditionally been a fundamental area of research for economists. A theory of the term structure is an explanation of the difference between the prices of risk free bonds with different maturity dates. The modern theories are based on two different sets of assumptions: no arbitrage or general equilibrium. It should be noted that apart from these differences both kinds of explanations are developed in the same framework: continuous time, diffusion processes for the state variables, frictionless and complete markets. In the no arbitrage approach, the price of a derivative asset based on interest rates is given by a replication argument: any two assets (or, more generally, any two investment strategies) which are equivalent in terms of future cash flows in every possible state of nature must have the same price at the current date. Harrison and Kreps [17] have shown that the assumptions of absence of arbitrage opportunities and of complete markets together imply the existence and the unicity of an equivalent probability measure (the so called "martingale probability", or "risk neutral probability") under which the price of every asset is equal to the expectation of the discounted future uncertain cash flows. This risk neutral probability is completely characterized by an adapted measurable process $\{\lambda_t\}$ called "risk premium process". The main drawback of the first generation of arbitrage pricing models is that the risk premium process has to be assumed, instead of being endogeneously derived from more fundamental assumptions about the structure of the economy. Examples of arbitrage pricing models of this kind are Vasicek [25] and Brennan and Schwartz [3]. In a second generation of arbitrage free models, which has been developed recently, the term structure is modelled using diffusion processes with time varying parameters, which both allow to avoid the specification of a risk premium process and to fit exactly the current term structure. Examples of recent arbitrage pricing models are Heath, Jarrow and Morton [19] and Hull and White [20].

The equilibrium approach requires instead a complete description of the underlying economy: the utility function of the representative agent, the technology of the production processes, the investment opportunities. Given these assumptions, it is possible to endogeneously derive the relation between the riskless instantaneous interest rate and the forcing variables (the

factors), together with the exact form of the risk premium. The most famous equilibrium model, which has also been intensively studied in a number of countries, is the one of Cox, Ingersoll and Ross [9] (henceforth CIR).

A common result of the great majority of these models is an exponential-affine formula for the price of zero coupon bonds. Let Y_1,Y_2,...,Y_K be the K factors driving the dynamic evolution of the term structure over time. These models frequently provide a valuation formula for zero coupon bonds of the following type:

$$P(t,T;\mathbf{Y}_t,\beta) = \exp\left\{A(t,T,\beta) + \sum_{h=1}^{K} B_h(t,T,\beta)Y_{h,t}\right\} \tag{1}$$

where t is the current date, T is the bond maturity date, the functions $A(t,T,\beta)$ and $B_h(t,T,\beta)$ depend on the parameters of the model (those appearing in the stochastic processes of the forcing variables and of the market prices of risk, one for each factor, collected in the vector β), and $\mathbf{Y}_t$ is the current realization of the factors. Duffie and Kan [11] have shown that valuation formulae like (1) are immediately obtained as soon as affine drifts and volatilities are assumed for the diffusion processes of the state variables, and the instantaneous interest rate is an affine function of these variables.

A notable property of formula (1) (but which is also shared by any valuation formula derived in a complete market framework) is that the price of the derivative asset is given *exactly*. In this sense, (1) can be interpreted as an accounting relationship between the price of the asset and the current realization of the state variables. When the number of assets is greater than the number of state variables, and given some value for β, a number of deterministic relationships arise between the prices of zero coupons at the same date but with different maturity dates. Strictly speaking, these models exclude the possibility of a discrepancy between observed and theoretical prices. Most empirical analysis of these models are however (explicitly or implicitly) based on the inclusion in the RHS of (1) of a random error.

In some cases the factors are identified with some known economic variables such as the instantaneous interest rate, an yield-to-maturity, or other macroeconomic variables. Nevertheless, the observation of these variables often presents some difficulties: macroeconomic aggregates are usually observed every month (if not every quarter), whereas financial variables are available at much higher frequency. Moreover, measurement errors and issues of definition of the same aggregates cannot be excluded. In some cases

it is necessary to recover to a proxy of the otherwise unobservable state variable, e.g. using a short yield to maturity or an overnight rate instead of the instantaneous interest rate. In any case, the use of a noisy observation, or of a proxy of the variable of interest, is likely to generate some effect in the inference stage whose properties are largely ignored. In this paper I suggest to interpret the forcing variables as latent, and to make inference on the parameters of the model using observations on bond prices and yields, which form an almost ideal dataset. The estimation and test of these continuous time, multivariate, dynamic latent variable model is greatly simplified by using some recently developed econometric techniques and by exploiting the particular structure of formulae like (1).

Exponential-affine formulae for the prices (or equivalently affine formulae for the yields to maturity) of zero coupon bonds are a maintained assumption in this paper, independently from the model originating them. In the following I concentrate for ease of exposition on the CIR [9] equilibrium model, but the estimation and testing approach that I suggest can be easily adapted to other frameworks. The rest of the paper is organized as follows. In section 2 I present a survey of the main approaches which have been adopted to estimate and test continuous time models of the term structure of interest rates. Section 3 presents the CIR multifactor model of the nominal term structure and discusses its "extension" based on the inclusion of random measurement errors. In section 4 I shall introduce a multivariate approach to statistical inference in these models under the assumption that it is possible to observe the prices of zero coupon bonds for a large number of maturity dates. This assumption is clearly not always acceptable, but it does nonetheless allow a simple exposition of the general principle of the multivariate approach. Section 5 concludes.

2 A survey of empirical work on the term structure in continuous time

It is not surprising that modern theories of the term structure have attracted a lot of attention. These models have been applied to a variety of different context, and a number of statistical procedures have been suggested to estimate and test them. In this section I focus on the approaches which explicitly deal with the problem of estimating the parameters of the model or of some

extension of it. I do not consider the empirical work whose objective is to test some implication of the models without fully estimating the parameters.

It is commonly observed that there exists a discrepancy between the degree of sophistication used in the derivation of theoretical models and the one characterizing their empirical implementations. In some cases the discrepancy simply refer to the difficulty to work with a model which provides *exact* valuation formulae, thus excluding the possibility to exploit the standard econometric tool kit. In this respect, some of the approaches that I outline in the following implicitly assume a particular structure of observation errors, thus violating the implications of the model. Some other approaches respect the feature of the formulae, but in this case different problems arise, related either to the properties of the variables approximating the factors, or to the quantity of information exploited to estimate the model.

Brown and Dybvig [4] first proposed a cross section procedure to simultaneously estimate the parameters and the state variables in the CIR [9] one factor model[1]. Their idea is extremely simple: let $P^*(t,T_i)$, $i=1,2,...,M$ be the observed bond prices at date t for M different maturity dates (bonds can be with or without coupon), and let $P(t,T_i;\mathbf{Y}_t,\beta)$ be the theoretical prices implied by a given model. At each date, it is possible to estimate β and the realization of $\mathbf{Y}_t$ (assumed unobservable) by minimizing the (weighted) sum of squared differences between $P^*(t,T_i)$ and its theoretical counterpart. It should be noted that this approach implicitly assumes the existence of normally distributed errors on the RHS of (1), independently distributed over time. Moreover, as it is only based on bond prices, it does not allow a separate identification of every parameter, but only of a limited number of functions of elements of β. It has been noticed that this procedure produces a very good fit of the model to the data, and that this result is not surprising since we estimate a different β at each date. By ignoring the structural relationship between β and the dynamic evolution of $\mathbf{Y}_t$, we implicitly augment (sometimes remarkably so) the number of factors.

A second approach assumes the observability of the state variables, and exploits the properties of their distribution. Chan, Karoly, Longstaff and Sanders [5], for example, assume the observability of the instantaneous riskless rate and estimate by GMM a variety of diffusion models for it. Prices of derivatives assets could then be computed by numerical techniques (such

[1]It should however be noticed that this procedure can be applied with minor modifications to a multifactor model.

as Monte Carlo integration or finite difference methods), once a particular form for the risk premium has been (separately) estimated or otherwise assumed. Aït-Sahalia [1] estimates nonparametrically the diffusion coefficient of a mean reverting process for the same rate, and compares the theoretical bond prices from a CIR [9] model with those arising from his generalized model. Longstaff and Schwartz [22] consider a two factor equilibrium model, where the dynamics of the yield curve is driven by the instantaneous interest rate and its volatility. They assume the observability of the first, and use a GARCH model to get a filtered series of the second factor. From a general viewpoint, this approach presents some problems, related to the use of proxies of the true state variables and to the need of separately estimate the risk premium parameter in order to empirically implement the model to derivative prices. Pearson and Sun [23] and Bianchi, Cesari and Panattoni [2] provide some interesting evidence about the consequences of the approximation error. A more subtle problem of this procedure, which has also been studied by Bianchi, Cesari and Panattoni [2], has to do with the stochastic properties of the time series of interest rates. Most of the models for this variable present a mean reverting drift term, which basically implies a first order autoregressive conditional expectation in discrete time. However, interest rates appear to be highly autocorrelated, and this means that estimates of the "mean reverting parameter" (which measures the speed of reversal to the "natural" or "long period" level of the rate) can be dramatically biased upwards even in fairly long samples[2]. Precise estimates of the parameters can be obtained by increasing the span of time covered by the sample, but this fact in turn introduces the possibility of including one or more structural breaks.

This time series approach has been recently refined by Pearson and Sun [23] in the context of the CIR [9] real two factor model. The general principle is however outlined in Duan [10], and can be described as follows. Consider a two factor model, which consists of (i) the conditional distribution of the (supposed unobservable) state variables, which depends on some unknown parameters, and (ii) some valuation formulae for particular (observed) derivative assets, depending on the same parameters *and* the risk premium. It is possible to exploit the deterministic one-to-one feature of the valuation rela-

[2]The estimation of the autocorrelation coefficient in discrete time autoregressive models presents the same difficulty. It is well known that the bias is a decreasing function of the sample length and an increasing function of the true unknown value of the parameter.

tionship to "invert" it and express the loglikelihood of the factors in terms of the observable prices. The new pdf will in general allow separate identification of all the parameters. Moreover, the inverted pricing formulae can be used to build a series of filtered realizations of the state variables, once an estimate of the parameters is obtained. Of course, this will only be possible if we limit ourselves to the observation of a number of assets exactly equal to the number of factors. Notice that this procedure does not violate the deterministic nature of the pricing formulae implied by the model: on the contrary, it is based on it. However, as witnessed by Pearson and Sun [23], it is likely that the constraint "number of derivatives = number of factors" limits the empirical identification of the parameters in some cases, since only a limited amount of the actually available information is taken into account. For example, the small sample bias discussed previously is likely to be a characteristic of this approach also.

To deal with this problem a solution has been suggested, which is once more based on the introduction of stochastic residuals in the RHS of formulae like (1). Chen and Scott [6] estimate three specifications of the CIR [9] nominal model, with one, two and three factors. The dataset consists of four time series of derivatives: two bills and two bonds. In every case, they use the inversion technique to transform some derivative prices in a realization of the latent state variables; since the number of assets is greater than the number of factors, they also introduce random residuals to reconcile theoretical and observed prices. The identification problems appear to vanish, at least for the one and two factors model, but their approach is limited, in the sense that to avoid undesired analytical complications they constrain themselves to the case "number of assets = number of state variables + number of random residuals". In other words, they suppose that *some* prespecified prices are observed without error, whereas some others are observed with error. The choice of the prices observed exactly is essentially arbitrary, apart from some intuitive considerations based on liquidity argument. A further problem with their procedure refers to the particular specification of the measurement errors they adopt. They suppose that the residuals are AR(1), and the parameters they estimate include both those of the term structure model and those of the error processes. This fact has two implications. First, it is hard not to see these autocorrelated residuals as a sort of "added" factors (of a very particular kind), since they contribute in the same way as the true factors to the cross and autocorrelation properties of the multivariate time series of yields. For example, a two (structural) factors model with two

autocorrelated residuals is in fact a particular specification of a four factors model. Since the total number of true and spurious factors is constant in every model they analyse, it is hard to use their procedure to find out the appropriate number of factors for a particular sample. Second, the particular structure of the error terms implies that the one factor model is not nested in the two factors one, and the latter is not nested in the three factors model. This complicates a formal statistical comparison between the models, even if the theoretical model naturally lends itself to the purpose.

Apart from the unusual specification of the error terms, the procedure supposes that some yields are observed exactly, whereas some others are not. It seems much more logical to suppose that *every* yield is subject to measurement errors of some kind, and to let the data decide the quality of the pricing model at every maturity. This means that the factors are now completely latent, and can not be filtered out of the yields by exploiting the deterministic feature of formulae like (1).

In a recent work, Chen and Scott [7] suggest to exploit the exponential affine structure of the zero coupon bond prices to estimate the model by applying the Kalman filter on the yields-to-maturity. This procedure, which is the first considered in this paper, is particularly well suited when prices on zero coupon bonds are available at many maturities. A more flexible one, based on the indirect inference approach, will also be introduced in the following.

3 The extension of the general CIR multifactor model

Although the framework I consider in this paper is suited to any continuous time model of the term structure which provides exponential affine formulae for zero coupon prices, in the following I shall focus for ease of exposition of the classical CIR [9] nominal multifactor equilibrium model of the term structure of interest rates (equations (57) - (60) in their paper). Given its general equilibrium foundations, this model allows the endogenous derivation of the risk premium which guarantees the absence of arbitrage opportunities. The assumed distribution of the state variables also implies the non negativity of the nominal yields at any maturity. The major drawbacks of the model are that it doesn't let to fit exactly the initial term structure, and it does provide

only quasi-explicit formulae of derivative assets involving option features, such as options of coupon bonds or on futures on coupon bonds, or futures involving a quality option. As this framework is well known, I shall give only a brief outline of it. CIR [9] assumptions include: log utility of the representative agent, square root processes for the state variables, and means and variances of the underlying production processes determined by linear functions of the sum of K state variables. The instantaneous nominal interest rate is assumed to be the sum of the state variables:

$$i = \sum_{i=1}^{K} y_i$$

and the square root processes for the state variables are:

$$dy_i = k_i(\theta_i - y_i)dt + \sigma_i\sqrt{y_i}dZ_i \tag{2}$$

for $i = 1, 2, ..., K$, where k_i, θ_i and σ_i are positive parameters, and $\{Z_i\}$ is a standard brownian motion defined on a probability space $(\Omega, \Im, P)$.

The derivation of the equilibrium of this economy provides the exact form of the risk premium, which is shown to be proportional to $\sqrt{y_i}$ for each state variable in (2). This result is determined by the covariability of the state variable with the marginal utility of nominal wealth. I shall define λ_i, $i = 1, 2, ..., K$, the K parameters characterizing the risk premium. It can be shown through the Girsanov Theorem that the dynamics of the state variables under the risk neutral probability Q (which is equivalent to the assumption of absence of arbitrage opportunities) is given by:

$$dy_i = [k_i\theta_i - (k_i + \lambda_i)\, y_i]\, dt + \sigma_i\sqrt{y_i}dZ_i.$$

The solution for the nominal price at time t of a nominally risk free zero coupon bond which pays 1\$ at time T is given by:

$$N(t, T; \mathbf{Y}_t, \beta) = \left\{\prod_{i=1}^{K} A_i(T - t; \beta)\right\} \exp\left\{-\sum_{i=1}^{K} B_i(T - t; \beta) y_{i,t}\right\} \tag{3}$$

which is, as expected, of the exponential affine form.

The yields to maturity are affine functions of the unobservable state variables:

$$R(t,T;\mathbf{Y}_t,\beta) = -\frac{1}{T-t}\ln N(t,T;\mathbf{Y}_t,\beta) \tag{4}$$
$$= -\frac{1}{T-t}\sum_{i=1}^{K}\ln A_i(T-t;\beta) + \frac{1}{T-t}\sum_{i=1}^{K}B_i(T-t;\beta)\,y_{i,t}$$

where:

$$A_i(T-t;\beta) = \left\{\frac{2\gamma_i e^{\frac{1}{2}(k_i+\lambda_i+\gamma_i)(T-t)}}{2\gamma_i + (k_i+\lambda_i+\gamma_i)\left[e^{\gamma_i(T-t)}-1\right]}\right\}^{\frac{2k_i\theta_i}{\sigma_i^2}},$$

$$B_i(T-t;\beta) = \frac{2\left[e^{\gamma_i(T-t)}-1\right]}{2\gamma_i + (k_i+\lambda_i+\gamma_i)\left[e^{\gamma_i(T-t)}-1\right]},$$

$$\gamma_i = \sqrt{(k_i+\lambda_i)^2 + 2\sigma_i^2},$$

and $\beta = (k_1, k_2, ..., k_K, \theta_1, \theta_2, ..., \theta_K, \sigma_1, \sigma_2, ..., \sigma_K, \lambda_1, \lambda_2, ..., \lambda_K)'$.

In this formulation, the econometric implications of the model are rather poor. Notice that the bond prices are deterministic functions of the realizations of the state variables $\mathbf{Y}_t$ and of the unknown parameters. The model just provides the conditional pdf for each state variable, which is known to be a noncentral χ^2 distribution with q_i degrees of freedom and parameters of noncentrality c_i :

$$f(y_{h,t} \mid y_{h,t-\Delta t}; k_i, \theta_i, \sigma_i) = \tag{5}$$

$$c_i e^{-c_i y_{h,t} - c_i e^{-k_i\Delta t} y_{h,t-\Delta t}} \left(\frac{y_{h,t}}{e^{-k_i\Delta t} y_{h,t-\Delta t}}\right)^{\frac{1}{2}q_i} I_{q_i}\left(2c_i\sqrt{y_{h,t}e^{-k_i\Delta t}y_{h,t-\Delta t}}\right)$$

where:

$$c_i = \frac{2k_i}{\sigma_i^2\left(1-e^{-k_i\Delta t}\right)},$$

$$q_i = \frac{2k_i\theta_i}{\sigma_i^2} - 1.$$

Δt is the length of the interval between two consecutive discrete time observations, and $I_q(.)$ is the modified Bessel function of the first kind of order q.

There are two ways of exploiting (5). The first is to assume that the state variables (or some one-to-one function of them) are observable, and to develop a standard maximum likelihood estimator of the parameters. As already stated in the introduction, this approach does not allow the estimation of the risk premium parameter, which can only be identified using bond prices. Moreover, the assumption of observable state variables is hardly acceptable for the reasons I stated in the introduction.

The second approach consists in assuming that at each date K bond prices are observed, and to exploit the deterministic relationships between these prices and the state variables to apply the MLE on the K observable prices instead of the K unobservable state variables. Notice that in this approach the K factor model can be estimated using *only* K bond prices. If more than K prices were observed, we could deduce some deterministic relations between them, and their joint distribution would have to be degenerate. If less than K prices were observed, we would not dispose of enough observations to solve the nonlinear system (4) with respect to $\mathbf{Y}_t$, and MLE would be unfeasible.

The second approach allows the estimation of the risk premium parameters λ_i, but it can nonetheless be criticized on the ground of the extremely limited information it is based on. Indeed, a one factor model *must* be estimated using a single point on the yield curve at each date. If the purpose of the analysis is just to recover the parameters, then this constraint is not troubling. However, if the objective is to evaluate comprehensively the quality of the fit of the model to the data, the constraint is much more puzzling. Moreover, if we stick to the assumption of observing at every date the *same* point on the yield curve, it is clear that the conclusions to be drawn could be considerably different using different points (see e.g. Pearson and Sun [23]). Some results could be discarded in favour of some others because they don't look enough "reasonable", but still room is left to the possibility of having different, and equally plausible, set of estimates.

The main reason of this apparent paradox is the deterministic feature of equations (4). Indeed, the approach just mentioned makes the implicit assumption that K prices are observed without error, while a random residual term is added on the RHS of (4) to allow an assessment of the fit of the theoretical yield curve to the observed one.

Strictly speaking, the inclusion of these random "measurement" errors totally invalidates the model. Indeed, it is clear that any (however small) difference between the observed and the theoretical price of an asset gives rise to an arbitrage opportunity, which means that the model is false and useless. However, under such a strict interpretation, *any* standard continuous time pricing model is false, because the immediate consequence of frictionless and complete markets is some deterministic relation like (4), which is immediately rejected by the data[3]. Notice that random residual terms are allowed in models based on an incomplete market hypothesis, as in Clement, Gourieroux and Monfort [8]. In this paper, however, I stick to a complete market framework.

It is hard to see how to get an empirical assessment of these models without the inclusion of measurement errors in (4). The strict interpretation stated above would be justified if the observations we dispose of would exactly meet the definitions of the model. Real data are however flawed by a number of actual measurement errors and of market imperfections, such as transaction costs, bid-ask spreads, nonsyncronous quotations, quotation errors, and so on. Under this point of view, the inclusion of error terms, even if by no means innocuous, looks much more reasonable.

Allowing random errors on the RHS of (4) leads us to two important remarks. First, it is hard to see why some bond prices should be observed exactly, while some others should not. Notice that this is exactly the case of the approach outlined earlier, where K prices are observed without errors, and every other point on the yield curve is subject to measurement error. A much more reasonable assumption would be to let each yields be observed with error, and to let the scale of the errors (their variance) to be estimated jointly with the other parameters of the model.

Second, if errors have to be included, their statistical properties must be completely specified. While it is hard to get a precise idea of their distribution, a zero mean assumption seems reasonable. Further properties can be derived from the definition of "factors" we adopt. If we define the factors as some variables governing the dynamics of bond prices, then measurement errors should not be autocorrelated, since otherwise they would appear as a kind of factors themselves. If we further define the factors as some variables which can also completely explain the contemporaneous correlation between

[3]The only model which could be compatible with the observation of M point on the yield curve is a M factor model.

any two bond yields, then measurement errors should also be uncorrelated one with each other. While the first point seems hardly questionable, in this paper I shall consider serially *and* contemporaneously uncorrelated error terms.

Finally, I assume a normal distribution for the residual terms. The "extended" CIR model is then formed by the following set of equations:

$$dy_i = k_i(\theta_i - y_i)\,dt + \sigma_i\sqrt{y_i}dZ_i, \qquad i = 1, 2, ..., K, \quad Z_i \perp Z_j \tag{6}$$

$$R(t, T_m; \mathbf{Y}_t, \beta) =$$

$$-\frac{1}{T_m - t}\sum_{h=1}^{K} \ln A_h(T_m - t; \beta) + \frac{1}{T_m - t}\sum_{h=1}^{K} B_h(T_m - t)\, y_{h,t} + \epsilon_{m,t}$$

where $\epsilon_{m,t} \sim NID(0, \omega_{mm})$, $m = 1, 2, ..., M$ denotes the maturities of the bonds whose price is observed, and A_i and B_i are as previously defined[4].

Finally, it is apparent how this approach can cope with the purposes of this paper. As stated in the introduction, the information in the whole term structure at different dates should be used to estimate the model, in order to avoid to work with samples covering long intervals of time, during which the occurrence of structural breaks is highly probable. In other words, instead of estimating the parameters by supposing the observability of the state variables, and working with a long sample of the factors, I wish to use short samples of a large number of yields to maturity, which are known functions of the (assumed unobserved) factors.

4 Statistical inference using zero coupon bond prices

In this section I shall assume that it is possible to observe the price of zero coupon bonds at every maturity. Even if in practice this is not true, the

[4]The normality assumption is essentially made for ease of presentation. It is somewhat problematic, since it does not exclude the possibility of negative nominal yields, but the procedure described in the following can easily handle different distributions of the measurement error, such as a lognormal, provided that the first two conditional moments are correctly specified.

approaches that I outlined below are very simple, and could be fruitfully employed if one trusts some fitted day-by-day yield curves (e.g. through polynomial splines, nonparametric regression, or some other method - see Gourieroux and Scaillet [16]). It would be interesting to analyse and measure the estimation bias introduced by working on fitted term structures instead of observed ones.

4.1 Pseudo maximum likelihood estimation

Let me recall that the model to be estimated is given by (6). The conditional pdf of each y_i is a noncentral χ^2 with degrees of freedom and noncentrality parameters which are known functions of unknown structural parameters. For ease of exposition, let me focus on the case $K = 1$:

$$y_t = a_0 + b_0 y_{t-\Delta t} + c_{0,t-\Delta t}\eta_t \tag{7}$$

$$R_{m,t} = a_m + b_m y_t + \epsilon_{m,t} \qquad m = 1, 2, ..., M$$

where:

$$\mathbf{E}\left(y_t \mid y_{t-\Delta t}\right) = a_0 + b_0 y_{t-\Delta t}$$

$$a_0 = \theta\left(1 - e^{-k\Delta t}\right)$$

$$b_0 = e^{-k\Delta t}$$

and:

$$\mathbf{Var}\left(y_t \mid y_{t-\Delta t}\right) = c^2_{0,t-\Delta t} = y_{t-\Delta t}\frac{\sigma^2}{k}\left(e^{-k\Delta t} - e^{-2k\Delta t}\right) + \theta\frac{\sigma^2}{2k}\left(1 - e^{-k\Delta t}\right),$$

where η_t is (conditionally on $y_{t-\Delta t}$) a standardized noncentral χ^2, with zero mean and unit variance. Notice that this implies that η_t is *uncorrelated* with $y_{t-\Delta t}$ (and else with $\eta_{t-\Delta t}$), but it will not in general be *independent* from $y_{t-\Delta t}$, since its moments of order higher than two do depend on $y_{t-\Delta t}$. Finally:

$$a_m = -\frac{1}{T_m - t}\frac{2k\theta}{\sigma^2}\ln\frac{2\gamma e^{\frac{1}{2}(k+\lambda+\gamma)(T_m-t)}}{2\gamma + (k+\lambda+\gamma)\left[e^{\gamma(T_m-t)}-1\right]} \qquad m = 1, 2, ..., M$$

$$b_m = \frac{1}{T_m - t}\frac{2\left[e^{\gamma(T_m-t)}-1\right]}{2\gamma + (k+\lambda+\gamma)\left[e^{\gamma(T_m-t)}-1\right]} \qquad m = 1, 2, ..., M$$

$$\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{M,t})' \sim NID(0, \Omega)$$

$$\Omega = diag\left(\omega_{11}, \omega_{22}, ..., \omega_{MM}\right).$$

Apart from the non gaussian distribution of η_t, model (7) looks very much like a standard state space model, with a transition equation for the unobserved state variable y_t, and a system of M measurement equations for the observed variables (yields to maturity). There are three features of model (7), all of them collected in the transition equation, which prevent us from considering the system a standard state space model: (a) the heteroskedasticity of the noise term, which depends on the unknown variable $y_{t-\Delta t}$; (b) η_t is uncorrelated but not independent from $y_{t-\Delta t}$; and (c) η_t is not normal, but rather a standardized non central χ^2. Nevertheless, the quasi-state space formulation can be exploited by making use of the standard Kalman filter to recover a filtered series of the latent variable for a given value of the parameters, and to plug the filter into a pseudo maximum likelihood estimation of the parameters of the structural model. If we correctly specify the first two conditional moments of y_t given $y_{t-\Delta t}$, the estimator thus defined is consistent and asymptotically normal (see Gourieroux and Monfort [13]). The asymptotic covariance matrix is not however given by the standard Fisher information matrix, but instead by the following product:

$$\left[J(\beta^0)^{-1} I(\beta^0) J(\beta^0)^{-1}\right]^{-1} \tag{8}$$

where:

$$I(\beta^0) = \mathbf{E}\left[\frac{\partial \ln L(\beta^0)}{\partial \beta}\frac{\partial \ln L(\beta^0)}{\partial \beta'}\right]$$

and:

$$J(\beta^0) = \mathbf{E}\left[\frac{\partial^2 \ln L(\beta^0)}{\partial\beta\partial\beta'}\right].$$

If η_t were normal (as it would be the case in a multivariate Vasicek type model), the Kalman filter would yield minimum square estimates (MMSEs) of the state[5]. Since η_t is not gaussian, I suggest to proceed *as if* the model were conditionally gaussian. Following Harvey, Ruiz and Sentana [18], the Kalman filter can be defined as *quasi optimal*.

The generalized model can easily be tested. The goodness of fit of models with a different number of factors can de directly compared through a likelihood ratio test, because the smaller model is nested in the larger one. If the state space model is not gaussian, the asymptotic distribution of the test statistic is not the standard χ^2 one, but it is rather a weighted sum of χ^2, as described in Vuong [26][6].

Several remarks are in order here. First, following Kim and Shephard [21], it would be possible to construct a modified algorithm (based on a mixture of normals to approximate the pdf of η_t, and simulation filtering techniques) to deal *exactly* with the non gaussianity of η_t. However the resulting filter would be numerically cumbersome, and I prefer to follow another approach, based on indirect inference, to cope with the small sample bias arising from the quasi optimality of the standard algorithm in this non gaussian context.

Second, apart from the task of estimating the parameters, any practical implementation of models of the term structure is likely to filter out of the observed sample a time series of observed factors. The Kalman filter is very well suited to this task, as it naturally provides smoothed estimates of the within sample realizations of the state variables, as well as forecasts of their future values. It thus becomes possible to use the filter in order to predict what the term structure will look like at any future date.

Third, for ease of exposition the presentation has focused on the one factor case. It is however clear that no new problems would arise in a multifactor model. Notice that in this last case the noises in the transition equations are

[5]Also, if the generalized model was gaussian (as it would be the case in the Vasicek [25] model), the ìn the state space approach would provide maximum likelihood estimates of the parameters. In this case, the asymptotic covariance matrix of the parameters given in (8) simplifies to $I(\beta^0)^{-1}$, since $I(\beta^0) = J(\beta^0)$.

[6]Again, if the state space model is gaussian, the asymptotic distribution of the test statistic is the standard one.

mutually independent, as they are originated by K independent brownian motions.

4.2 Indirect inference estimation

The estimator based on the Kalman filter seems appealing for a number of reasons, but it still leaves room to some criticism. Its "quasi optimality" refers to the non gaussian feature of the transition equation. A failure to take into account the exact distribution of the state variables implies the loss of the asymptotic efficiency, as demonstrated by the asymptotic covariance matrix (8). However, notice that the small sample properties look more important in the context of samples covering small intervals of time and a large number of bonds that I consider in this paper. It is well known that even MLE can be biased (and sometimes very much so) in estimating parameters in conditional models on short samples of autocorrelated observations. As an example, it is well known that MLE of the autoregressive parameter in a standard AR(1) gaussian model is biased downwards, and that the bias is an increasing function of the true value of the parameter, and a decreasing function of the sample size.

In the context of continuous time model with mean reverting drift (such as the square root ones assumed for the state variables), it is well known that the small sample bias mainly shows up in the parameter k. As Bianchi, Cesari and Panattoni [2] have noticed, it seems hard to obtain *unbiased* estimates of k in small samples using direct procedures.

This is the reason which underlies the interest that "indirect" procedures have met in this context; see Smith [24], Gourieroux, Monfort and Renault [14] (GMR in the following) and Gallant and Tauchen [12] for a general presentation. It should be noticed that the issue of the estimation of the autocorrelation parameter is one of the themes of Gourieroux, Renault and Touzi [15], which provides encouraging results of the relative performance of the indirect estimator with respect to the traditional OLS one. They also prove that when the ratio of the length of the simulated series with respect to the observed one diverges ($H \to \infty$ in the notation of GMR), the indirect estimator automatically operates a second order bias correction in small samples, which is a known property of Bootstrap estimation procedures.

The auxiliary model that I consider is based on the Kalman filter. The procedure thus provides a "corrected" estimator, which should act as a reduction of the bias which characterizes the small sample behaviour of the

otherwise consistent Kalman approach. Let denote with β the whole vector of parameters (those appearing in the structural term structure model and those characterizing the dispersion of the measurement errors). The whole procedure can be outlined as follows:

1. Estimate the model using the quasi optimal Kalman filter on real data, thus obtaining $\widetilde{\beta}_T^K$; notice that a distinction should be made between the structural parameters and the parameters of the auxiliary model only in the case of differences between the two. In the approach outlined here, however, the two models coincide, and the only disparity is between estimation methods. Therefore I shall denote with $\widetilde{\beta}_T^K$ the estimates based on the Kalman filter, and with $\widetilde{\beta}_T^I$ those based on the indirect procedure.

2. Calibrate model (7) (or a K-factor generalization of it) by picking up the values of the parameters that, when used in a simulation of (7), can best account of $\widetilde{\beta}_T^K$.

Mathematically, the second step is equivalent to solving either one of the three following optimization problems:

$$\widetilde{\beta}_T^{I_1} = \arg\max_{\beta} Q_T\left[\underline{x}_T^1, \widetilde{\beta}_{HT}^K(\beta)\right] \qquad \text{(see Smith [24])}$$

$$\widetilde{\beta}_T^{I_2} = \arg\min_{\beta} \left\|\widetilde{\beta}_T^K - \widetilde{\beta}_{HT}^K(\beta)\right\|_{\Omega} \qquad \text{(see GMR [14])}$$

$$\widetilde{\beta}_T^{I_3} = \arg\min_{\beta} \left\|\frac{\partial Q_T}{\partial \beta}\left[\widetilde{\underline{x}}_{HT}^1(\beta), \widetilde{\beta}_T^K\right]\right\|_{\Omega} \qquad \text{(see Gallant and Tauchen [12])}$$

where: $\underline{x}_T^1$ is the observed sample, of length T; $\widetilde{\underline{x}}_{HT}^1(\beta)$ is a simulated sample drawn from (7) for a given value of the parameters and of length HT; $Q_T(x, \beta)$ is the objective function of the auxiliary estimation criterion (the Kalman filter procedure); $\widetilde{\beta}_{HT}^K(\beta)$ is the estimate of β obtained by maximizing Q on $\widetilde{\underline{x}}_{HT}^1(\beta)$. The relations between the three estimators are the following: when the number of parameters is the same in the auxiliary and the structural model ("exact identification", as it is the case here), the three coincide exactly. When the auxiliary model has more parameters than

the structural one ("over identification"), $\widetilde{\beta}_T^{I_2}$ and $\widetilde{\beta}_T^{I_3}$ are asymptotically equivalent, and they have a smaller covariance matrix than $\widetilde{\beta}_T^{I_1}$. It is also possible to state that the former two are not as efficient as MLE, unless the auxiliary model "smoothly embeds" (in Gallant and Tauchen [12] terms; basically, it is a reparameterization of) the structural model, and $H \to \infty$.

Some remarks can be made at this point. First, as the structural and the auxiliary model coincide, so do the two vectors of parameters. Therefore the auxiliary parameters can be most easily interpreted under the structural model. This fact provides the intuition for a very efficient iterative algorithm for the solution of the second kind of optimization problem of indirect estimation, first advanced by Gourieroux, Monfort and Renault [15]. The algorithm is based on the fact that the function:

$$g_T(\beta) = \beta + \widetilde{\beta}_T^K - \widetilde{\beta}_{HT}^K(\beta)$$

is a strong contraction, with the indirect estimator $\widetilde{\beta}_T^{I_2}$ as its unique fixed point. Thus, for a given $\widetilde{\beta}_T^K$, it is possible to construct the sequence $\left\{\widetilde{\beta}_T^{I_2}(n)\right\}_{n\geq 0}$ such that:

$$\widetilde{\beta}_T^{I_2}(0) = \widetilde{\beta}_T^K \qquad \text{and} \qquad \widetilde{\beta}_T^{I_2}(n+1) = g_T\left[\widetilde{\beta}_T^{I_2}(n)\right],$$

which converges towards its unique fixed point $\widetilde{\beta}_T^{I_2}$. Monte Carlo experiments in the AR(2) gaussian model reported in Gourieroux, Renault and Touzi [15] confirm the above intuition, as the algorithm regularly converges in a very low number of iterations. In the Monte Carlo analysis on the one factor CIR model reported below, convergence was always achieved in no more than 11 iterations.

A second remark is about the exact identification implicit in the previous approach. It is clear that by choosing a larger auxiliary model one could introduce some overidentifying restrictions which are potentially helpful if more precise estimates are desired. In this paper I do not pursue this idea for two reasons. First, there exists a trade off between precision of the estimates and numerical simplicity, if nor feasibility, of the procedure. The introduction of overidentifying restrictions has some undesirable consequences, such as for example the difficulties in understanding when a global minimum of the objective function is attained (in the exact identification case the global minimum corresponds to a zero value of the objective function), or the impossibility to apply the previous algorithm. Moreover, to efficiently exploit

the overidentifying restrictions it is necessary to estimate the optimal weighting matrix Ω, and it has been shown by Smith [24] that the noise implicit in such an estimate significantly worsens the small sample performance of the "asymptotically efficient" estimators $\widetilde{\beta}_T^{I_2}$ and $\widetilde{\beta}_T^{I_3}$.

4.3 A Monte Carlo Experiment

To investigate the feasibility and the properties of the two procedures outlined before, I designed a Monte Carlo experiment for a one factor CIR model. The values of the parameters used are those estimated in Chen and Scott [6] on a sample of nine years of weekly observations of prices of US bills and bonds: $k^0 = 0.6248$, $\theta^0 = 0.09304$, $\sigma^0 = 0.1054$ and $\lambda^0 = -0.09235$. The factor process has been simulated using the exact discretization, i.e. by generating pseudo random non central χ^2 variates with appropriately chosen number of degrees of freedom and parameter of non centrality. On the basis of this simulated time series of the state variable I constructed the prices of the two zero coupon bonds with three months and one year to maturity, ad I added to the yields to maturity independent zero mean homoschedastic normal measurement errors with a standard error of 0.15%. I created 500 replications of samples made up of 200 daily observations, and on each one of them I estimated the vector of parameters using both the Kalman filter and the indirect approach.

Table 1 reports some summary statistics of the results of the experiment. ML stands for the ML approach as implemented by Chen and Scott [6]; recall that thi sprocedure is based on the assumption that one price (in our case, the 13 weeks bill) is observed exactly, whereas the others are not. KALMAN stands for the Kalman filter based estimation, and II for the indirect inference estimator.

Figures 1, 2 and 3 depict the nonparametric kernel estimate of the densities of the three estimators for each parameter. The continuous line is the density of the indirect estimator; the dashed line is the density of the Kalman filter estimator, and the dotted - dashed line is the density of Chen and Scott [6] estimator. The main problem with the ML procedure are apparent: the volatility parameter and the variance of the measurement error on the one year bond are significantly overestimated, and the other parameters (apart from θ) are not precisely estimated. The interesting comparison to be made is between the quasi-optimal Kalman filter procedure and the indirect inference one. Both of them seem to perform rather well, since the small sample

average and median of the estimates are close to the true values for every parameter. It is not surprising that the Kalman estimators are less dispersed that the indirect inference ones; the latter inevitably show up some simulation noise, whose consequencies are most evident when H is low (in this experiment H was set equal to 2). By choosing a higher value of H the gap between the small sample dispersions would be reduced, and probably the same result could be obtained by picking innovation realizations using some variance reduction techniques. Even if the experiment is very simple and perhaps somewhat irrealistic, the overall evidence seems promising.

5 Conclusions

This paper has dealt with the problem of estimating a large class of continuous time models of the term structure of interest rates, characterized by the exponential affine structure of the zero coupon bond prices. Most no arbitrage and equilibrium models belong to these class.

I have started with a survey of the empirical work on the term structure in continuous time. A common property of this kind of models is that *exact* bond prices are provided. Otherwise stated, the assumption of absence of arbitrage possibilities rules out the existence of observation errors. Strictly speaking, this means that the model implies exact relationships between the prices at the same date and at different dates. If the observed prices do not satisfy these constraints, the model is useless and false. However, in the real world there are many sources of measurement errors, and researchers on the subject have usually adopted a more pragmatic point of view. In most cases, the theoretical model is tested by checking the distance between the theoretical (given some parameters' estimates) and the observed prices. Such a procedure implicitly assumes the existence of measurement and observation errors.

In this paper I suggest to explicitly take into account these errors, by constructing a "generalized" theoretical model, which provides the *expected value* of bonds and other derivative assets prices. In such a framework it is no more necessary to assume that the factors are observable, nor that they must be proxied. If zero coupon prices are exponential affine in the unobservable state variables, yields to maturity are affine in the factors. The model then has a usual state space structure, and the Kalman filter can easily be applied. Thus, if zero coupon bond prices are observed at a sufficiently wide

range of maturities, the generalized model can be estimated by maximizing the gaussian likelihood. If both the transition equations and the measurement equations are gaussian, maximum likelihood estimates are obtained; if some departures from gaussianity are observed, pseudo maximum likelihood estimates are obtained.

This procedure is straightforward, but it is based on the availability of zero coupon prices at a wide range of maturities. Usually, zero coupon bonds exist only at the short end of the term structure. Moreover, as it is the case for most estimation procedures of dynamic time series models, the estimates can be severely biased in short samples. A second approach suggested in this paper, based on the indirect inference principle, avoids both problems by exploiting simulation methods. I suggest to use the state space model as the auxiliary criterion. This allows to give an immediate interpretation of the auxiliary parameters, and to estimate the structural parameters using a simplified iterative algorithm.

As an application, I consider a Monte Carlo experiment for the one factor nominal CIR model. The results are encouraging: even if the sample covers a short interval of time (only one year of daily observations), the parameters' estimates seem rather precise, and in any case superior to those generated by an alternative approach. The comparison between the state space and the indirect inference estimators suggest that the latter may be superior on average, but that it also show up a larger dispersion around the true value. It is suggested that the latter drawback could be (at least partially) eliminated by using longer simulated series or variance reduction techniques.

References

[1] Y. Ait-Sahalia. Nonparametric pricing of interest rate derivative securities. mimeo, 1994.

[2] C. Bianchi, R. Cesari, and L. Panattoni. Alternative estimators of the Cox, Ingersoll and Ross model of the term structure of interest rates: A monte carlo comparison. Technical Report 236, Banca d'Italia, Temi di Discussione, 1994.

[3] M.J. Brennan and E.S. Schwartz. A continuous time approach to the pricing of bonds. *Journal of Banking and Finance*, 3(2):133–55, 1979.

[4] S.J. Brown and P.H. Dybvig. The empirical implication of the Cox, Ingersoll, Ross theory of the term structure of interest rates. *Journal of Finance*, 41:617–30, 1986.

[5] K.C. Chan, G.A. Karoly, F.A. Longstaff, and A.S. Sanders. An empirical comparison of alternative models of the short term interest rate. *Journal of Finance*, 47:1209–27, 1992.

[6] R.-R. Chen and L. Scott. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *The Journal of Fixed Income*, pages 14–31, december 1993.

[7] R.-R. Chen and L. Scott. Multi-factor cox-ingersoll-ross models of the term structure: Estimates and tests from a kalman filter model. mimeo, University of Georgia, January 1995.

[8] E. Clement, C. Gourieroux, and A. Monfort. Prediction of contingent price measures. Technical Report 9332, CREST-ENSAE, Paris, 1993.

[9] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[10] J.-C. Duan. Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 4(2):155–167, 1994.

[11] D. Duffie and R. Kan. A yield - factor model of interest rates. Technical report, Graduate School of Business, Stanford University, 1993. mimeo.

[12] R.A. Gallant and G. Tauchen. Which moments to match? *Econometric Theory*, 1994. Forthcoming.

[13] C. Gourieroux and A. Monfort. *Statistique et Modeles Econometriques.* Economica, Paris, 1989.

[14] C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of Applied Econometrics*, (8):S85–S118, 1993.

[15] C. Gourieroux, E. Renault, and N. Touzi. Calibration by simulation and small sample bias correction. Technical report, CREST-ENSAE, Paris, 1994. mimeo.

[16] C. Gourieroux and O. Scaillet. Estimation of the term structure from bond data. Technical report, CREST-ENSAE, Paris, 1994. mimeo.

[17] M.J. Harrison and D.M. Kreps. Martingale and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, 1979.

[18] A. Harvey, E. Ruiz, and E. Sentana. Unobserved component time series models with arch disturbances. *Journal of Econometrics*, 52:129–157, 1992.

[19] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992.

[20] J. Hull and A. White. Pricing interest-rate derivative securities. *Review of Financial Studies*, 3(4):573–592, 1990.

[21] S. Kim and N. Shephard. Stochastic volatility: Likelihood inference and comparison with arch models. Technical report, Nuffield College, Oxford, 1994. mimeo.

[22] F.A. Longstaff and E.S. Schwartz. Interest rate volatility and the term structure: A two factor general equilibrium model. *Journal of Finance*, 47:1259–82, 1992.

[23] N.D. Pearson and T.-S. Sun. Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll and Ross model. *Journal of Finance*, 49(4):1279–304, 1994.

[24] A.A. Smith. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8:S63–S84, 1993.

[25] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–88, 1977.

[26] Q.H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57:307–33, march 1989.

A Appendix: The Results of the Monte Carlo Experiment

TABLE 1
SUMMARY STATISTICS OF THE MONTE CARLO EXPERIMENT
ON THE 1 FACTOR CIR MODEL

Summary Statistics on k ($k^0 = 0.6248$)

	ML	KALMAN	II
Mean	0.865	0.732	0.681
Median	0.820	0.702	0.618
Standard Error	0.323	0.236	0.318
Minimum	0.296	0.279	0.0438
Maximum	2.913	2.357	2.386
RMSE	0.402	0.259	0.323
Kurtosis	7.796	7.569	5.270

Summary Statistics on θ ($\theta^0 = 0.09304$)

	ML	KALMAN	II
Mean	0.0927	0.0871	0.103
Median	0.0891	0.0838	0.0927
Standard Error	0.0273	0.0254	0.0481
Minimum	0.0244	0.0242	0.0232
Maximum	0.209	0.195	0.353
RMSE	0.0273	0.0261	0.0491
Kurtosis	4.012	4.103	6.979

Summary Statistics on σ ($\sigma^0 = 0.1054$)

	ML	KALMAN	II
Mean	0.185	0.104	0.105
Median	0.182	0.105	0.105
Standard Error	0.0228	0.00960	0.0129
Minimum	0.132	0.0732	0.0646
Maximum	0.313	0.136	0.157
RMSE	0.0825	0.00963	0.0129
Kurtosis	6.116	3.254	3.272

Summary Statistics on λ ($\lambda^0 = -0.09235$)

	ML	KALMAN	II
Mean	-0.172	-0.193	-0.146
Median	-0.136	-0.164	-0.0975
Standard Error	0.240	0.222	0.286
Minimum	-2.075	-1.840	-1.909
Maximum	0.262	0.236	0.366
RMSE	0.253	0.244	0.291
Kurtosis	11.765	9.320	6.746

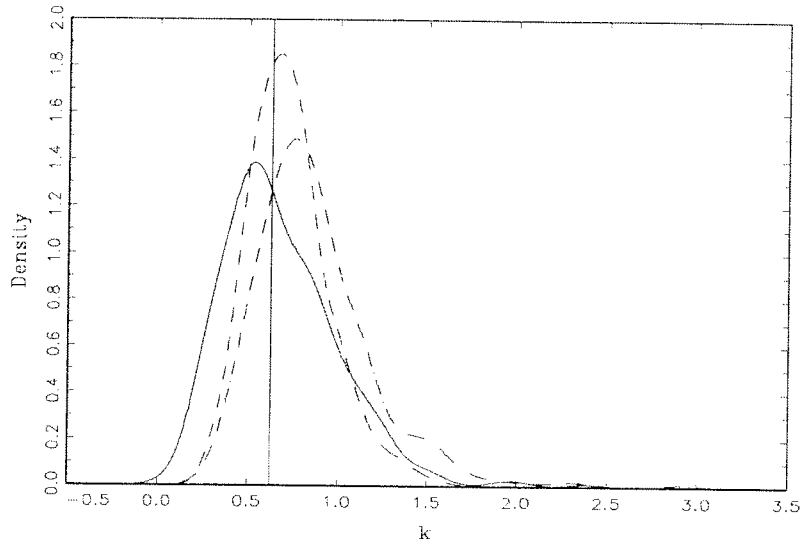
Summary Statistics on ω_{11} ($\omega_{11}^0 = 0.15$)

	KALMAN	II
Mean	0.149	0.150
Median	0.150	0.150
Standard Error	0.0118	0.0163
Minimum	0.116	0.107
Maximum	0.192	0.209
RMSE	0.0118	0.0163
Kurtosis	3.090	3.236

Summary Statistics on ω_{22} $(\omega_{22}^0 = 0.15)$

	ML	**KALMAN**	**II**
Mean	0.192	0.150	0.151
Median	0.192	0.150	0.150
Standard Error	0.0101	0.0103	0.0142
Minimum	0.158	0.119	0.105
Maximum	0.227	0.179	0.191
RMSE	0.0431	0.0103	0.0142
Kurtosis	3.093	2.965	2.899

FIGURE 1
KERNEL ESTIMATES OF THE DENSITIES OF 3 DIFFERENT ESTIMATORS
PARAMETERS k AND θ



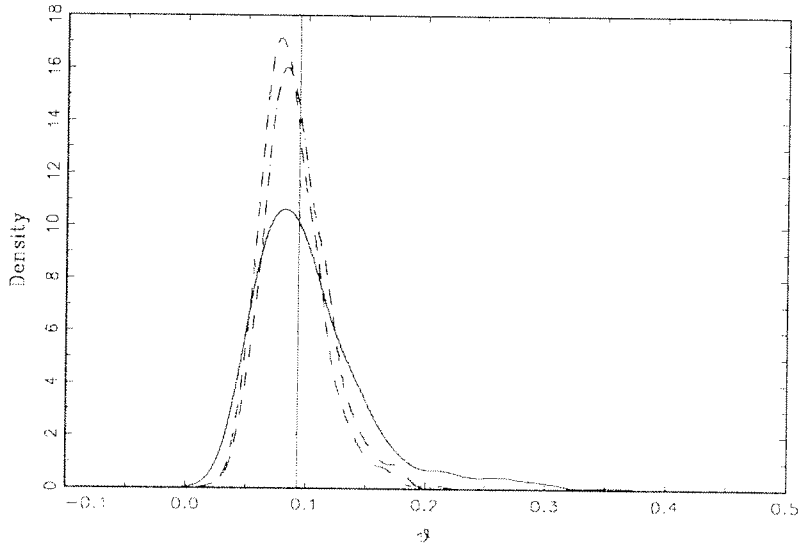
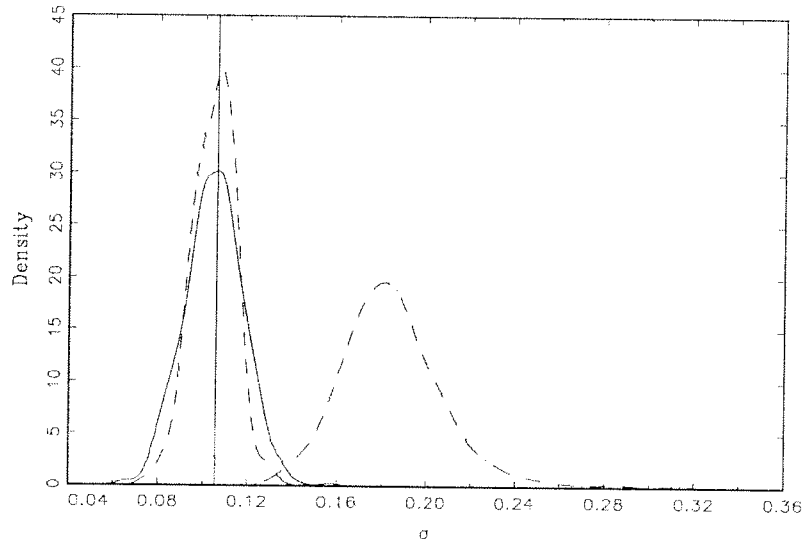


FIGURE 2
KERNEL ESTIMATES OF THE DENSITIES OF 3 DIFFERENT ESTIMATORS
PARAMETERS σ AND λ



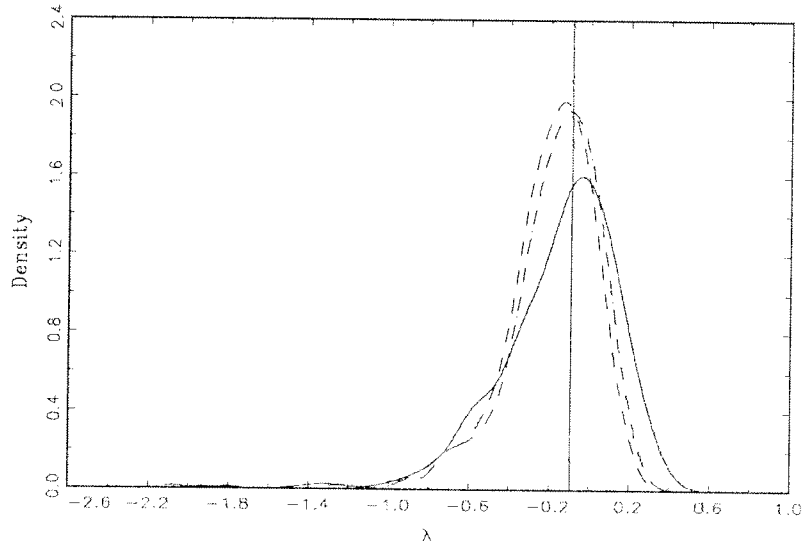


FIGURE 3
KERNEL ESTIMATES OF THE DENSITIES OF 3 DIFFERENT ESTIMATORS
PARAMETERS ω_{11} AND ω_{22}