ISSN 0265-8003

DIRECT ESTIMATION OF THE RISK NEUTRAL FACTOR DYNAMICS OF AFFINE TERM STRUCTURE MODELS

Dennis Bams and Peter C Schotman

Discussion Paper No. 2034
December 1998

Centre for Economic Policy Research
90–98 Goswell Rd
London EC1V 7DB
Tel: (44 171) 878 2900
Fax: (44 171) 878 2999
Email: cepr@cepr.org

CEPR Discussion Paper No. 2034

November 1998

ABSTRACT

Direct Estimation of the Risk Neutral Factor Dynamics of Affine Term Structure Models*

This paper proposes a panel data framework for tests of affine models of the term structure of interest rates which cover equilibrium (or endogenous) models as well as extended (or exogenous, evolutionary) models. The econometric model pools yield curve data for different moments in time. Since each cross-sectional yield curve only depends on the risk neutral factor dynamics, the estimator does not involve any assumptions on the price of risk, or on actual interest rate dynamics. In the empirical application one and two factor Gaussian models are tested on US interest rate data. The main empirical results are: (i) that a two-factor model cannot be rejected; (ii) that mean reversion is highly significant; and (iii) that the extended models are 'over-differenced'.

JEL Classification: C33, G13
Keywords: affine models, panel data, term structure of interest rates, factor models

Dennis Bams and Peter C Schotman
Maastricht University
LIFE
PO Box 616
6200 MD Maastricht
THE NETHERLANDS
Tel: (31 43) 388 3838/62
Fax: (31 43) 325 8530
Email: w.bams@berfin.unimaas.nl
p.schotman@berfin.unimaas.nl

*Dennis Bams acknowledges financial support from the Dutch Organization of Scientific Research (NWO). The authors would like to thank Frank DeJong, Christian Gouriéroux and Ronald VanDijk for fruitful discussions of an earlier version. They also thank various participants at the CEPR Finance meetings in Gerzensee, Econometric Society European meeting, European Finance Association meeting, Seventh Bi-annual Panel Data conference, Amsterdam, Finance and Econometric meeting, Tilburg, Stockholm Business School and HEC, Paris for their comments. All errors are the authors' own.

Submitted 29 October 1998

NON-TECHNICAL SUMMARY

Statistical models of the term structure of interest rates are often used to price interest rate derivative securities. For that purpose the models should satisfy a number of properties. First, since some derivative securities have complicated option features, computational considerations limit the number of stochastic factors that govern the dynamics of the entire yield curve. Second, the evolution of the term structure should be free of arbitrage opportunities. Third, the dynamics must be specified under the risk neutral measure instead of the actual probability measure. The affine class of term structure models developed in Duffie and Kan (1996) is very attractive, since it provides a tractable affine relation between yields of different maturities and the factors, and a dynamic structure from which it is straightforward to build a tree for option pricing applications.

This paper aims at direct estimation of the risk neutral factor dynamics. The risk neutral dynamics differ from the actual time series dynamics by a possibly time varying price of risk. The risk neutral parameters also determine the shape of the yield curve at every moment in time and can be estimated from a panel of observed yield curves. The advantage of this method is that it avoids the assumption that the price of risk is constant and is thus more robust than time series oriented estimators. Ignoring the time series information, however, can result in a loss of efficiency. In a Monte Carlo experiment we find that the cross-sectional estimator is highly efficient when yield curves that start at very different levels are pooled. For example, mean reversion can be estimated very precisely if we have observations on a yield curve with a short rate of 12% and another one that starts at 4%. When there is sufficient variation in short-term interest rates, the time series information does not add much to the pooled cross-sections.

For the econometric inference we develop a fixed effects panel data estimator for the risk neutral parameters of the model. The econometric model is suitable for both equilibrium as well as extended term structure models. The equilibrium models describe the shape of the yield curve, whereas the extended models describe the arbitrage free evolution of the yield curve from an exogenously given initial yield curve. In the application to a panel of 25 years of US interest rates we find that a two factor Gaussian model fits the factor structure of the data well. Both equilibrium and extended versions of the model are subjected to a number of diagnostic tests. The extended version model turns out to be over-differenced in the sense that the residuals are negatively autocorrelated. Overall, the equilibrium and extended models are very similar. For both models the parameter estimates point to one factor that is close to being non-stationary and a second factor with strong mean reversion.

Although consistent with time series characteristics, the near unit root behaviour of the first factor is an implication of the panel of cross-sectional yield curves. An equilibrium model always implies that the infinite horizon yield is a constant. Since ten year rates show a lot of variation in the data, the only way an equilibrium model can fit the data is having one almost non-stationary factor. Similarly, an extended model can only explain why changes in long-term rates are not much smaller than changes in short-term rates, if the dominant factor has very low mean reversion.

1 Introduction

Models of the term structure of interest rates provide the shape and dynamics of the yield curve from assumptions about the dynamics of some underlying factors and the price of the risk associated with each factor. A limited number of underlying state variables are assumed to account for the behaviour of interest rates of many different maturities. The models are often applied in pricing various interest rate derivative securities. These applications require an estimate of the risk adjusted (or risk neutral) dynamics of the factors.

The risk neutral process can be inferred from the observed term structure of interest rates, or from its change relative to an initial term structure. To estimate the parameters of the process one would like to have an analytically tractable model that relates the yields to the factors. A class of models that has received a lot of attention is the affine class, analysed in detail by Duffie and Kan (1996), Campbell, Lo and MacKinlay (1997, ch 11) and Frachot and Lesne (1993). It has a tractable linear structure and nests a number of wellknown equilibrium term structure models like the Vasicek (1977) and Cox, Ingersoll and Ross (1985, CIR) model, and their multifactor generalisations.

Even though the affine class is tractable, direct estimation of the risk neutral parameters has not been very succesful. Many studies have attempted to estimate parameters from cross sectional yield curve data. In these papers the structural parameters are re-estimated every time period using the bond prices at that moment. All of them report unstable and erratic parameter estimates.[1] The econometric problems could be due either to a lack of information in yield curve data for a single date, or to misspecification of the models.

The aim of this paper is to revisit the direct estimation of the risk neutral process using panel data techniques. We pool monthly yield curves for a period of twenty-five years to reduce the problem of low power. As in the cross sectional studies cited above, the parameters are estimated without explicit reference to the time series properties of interest rates. Econometrically, this leads to a panel model with time dependent fixed effects.

From the cross sectional estimation we obtain the implied dynamic process of interest rates that generates yield curves that closely fit the observed term structure.

[1] See for example Brown and Dybvig (1986), Brown and Schaefer (1994), Dahlquist and Svensson (1994), DeMunnik and Schotman (1994), Addolorato and Berardi (1994), and Sercu and Wu (1997).

But rather than modeling the levels of the yield curve, applications to derivative pricing emphasize models of the yield curve in deviation of last period's term structure.[2] These models provide an expression for the shape of the current term structure, conditional on an exogenously given initial term structure. This initial term structure need not, but could, be consistent with an equilibrium model. Some of these models, for example the extended Vasicek and extended CIR models, belong to the extended affine class as analysed in Frachot and Lesne (1993). Our panel data analysis with fixed effects is suited for both types of term structure models. The equilibrium and extended models will generally emphasize different moments of the data. But depending on the properties of measurement error in the empirical model, the two formulations are shown to be closely related, and can be compared in a common econometric framework.

Following Jacquier and Jarrow (1998), Brown and Dybvig (1986) and others, we add an error term to the theoretical model. The covariance matrix of the error term is modeled explicitly to take account of heteroskedasticity and the cross sectional correlation of yields. We tightly parameterise the cross sectional covariance structure, since we wish to include a broad range of maturities in the empirical model. Moreover, the specification explictly deals with residual autocorrelation.

We find that the efficiency gains of the panel are large. For the one- and two-factor models that we estimate we obtain parameter estimates with very low standard errors. The first factor has very slow mean reversion, but with a t-ratio of more than eight mean reversion is significantly different from zero. The average convexity of the yield curve provides a sharp point estimate of the implied volatility of the factors. We also find that the extended and equilibrium models give very similar point estimates for the parameters of interest.

Our fixed effects model contrasts with other panel data studies of affine term structure models.[3] In these studies the cross sectional yield curve model is estimated jointly with a time series process for the factors. Combining cross section and time series information exploits even more information in the data. But time series data are related to the actual dynamics, so that additional assumptions on the factor risk prices must be made. The usual assumptions are that the price of risk is either

[2] See Hull and White (1990) and Heath, Jarrow and Morton (1992) for a theoretical analysis. See Jarrow (1996) and Hull (1996) for a textbook treatment.

[3] See for example Chen and Scott (1993, 1995), Gibbons and Ramaswamy (1993), DeJong (1997), Geyer and Pichler (1997), Frühwirth-Schnatter and Geyer (1997), Babbs and Nowman (1998), Lund (1997), Kappi (1997), Duan and Simonato (1995), Gong and Remolona (1996).

constant, or proportional to (one of) the factors. For the fixed effects analysis we do not need any assumption on the actual time series properties of the factors. Only the risk neutral process matters. Still we also obtain sharp estimates of the parameters, in particular mean reversion. Empirically, 25 years of pooled cross sections dominate the time series information.

The panel data analysis also allows a detailed misspecification analysis. Misspecification here refers to the shape of the yield curve, in contrast to the dynamic structure of interest rate processes. When we find misspecification for the yield curve, this misspecification will be present a fortiori in the panel models that combine time series and cross sectional information. For two-factor Gaussian models we can not reject the cross sectional restrictions. The two factor model is very parsimonious, containing only five structural parameters that are sufficient to describe the variety of shapes encountered in a history of twenty five years of monthly term structures. In addition to the five structural parameters, each yield curve in a two factor model has two time dependent parameters, only relevant to that particular yield curve.

For those models that appear correctly specified in the cross sectional dimension, it is interesting to compare our results to the other panel literature by imposing assumptions on the price of risk and the time series behavior of the factors. In a broad sense, the cross sectional information is consistent with recent empirical time series studies. Andersen and Lund (1996) and Balduzzi, Das and Foresi (1998) argue that interest rate dynamics contains at least two factors. One of the factors is a slowly evolving mean, and the second factor represents quickly mean reverting movements around this mean.

The remainder of this paper is organized as follows. Section 2 introduces notation and provides a brief review of affine term structure models. Section 3 considers the special case of the (extended) Vasicek model in detail. In section 4 we illustrate the problems of cross sectional estimation, and explain why pooling different cross sections will lead to a very efficient estimator. Section 5 discusses the econometric panel data model. The final part of this section discusses the misspecification tests. Section 6 describes the data and stylized facts of the US term structure. Section 7 reports the empirical results. Section 8 concludes.

2 Affine Term Structure Models

Duffie and Kan (1996) consider equilibrium formulations of the affine yield curve models. Frachot and Lesne (1993) generalise this approach to deal with the extended specifications, that match an initial yield curve exactly.[4] The general idea underlying the affine term structure models is that at time t there exist a vector of K factors, Z_t, that govern the term structure movements. The drift and the diffusion process of the factors are affine in the factors. The affine class has gained popularity because it leads to a tractable solution for the term structure of interest rates that is itself an affine function of the factors.

The risk adjusted process for the underlying factors in an affine term structure model is specified as

$$dZ_t = (\Phi_t - \Gamma Z_t)\,dt + V(Z_t)dW_t\,, \tag{1}$$

where Z_t denotes the vector of K factors, W_t is a K-dimensional Brownian motion under the risk-adjusted probability measure, and Φ_t is a $(K \times 1)$ deterministic function. The time variation in Φ_t enables an exact fit of an observed initial yield curve and hence represents the no-arbitrage case. When Φ_t is a vector of constants we are in the equilibrium model case of Duffie and Kan (1996). Mean reversion is determined by the $(K \times K)$ matrix Γ, and $V_t(Z_t)$ is defined such that the covariance matrix is affine in the factors

$$\Psi_t = V_t(Z_t)V_t(Z_t)' = \begin{pmatrix} \alpha_{11}Z_t + \beta_{11} & \cdots & \alpha_{K1}Z_t + \beta_{K1} \\ \vdots & & \vdots \\ \alpha_{K1}Z_t + \beta_{K1} & \cdots & \alpha_{KK}Z_t + \beta_{KK} \end{pmatrix} \tag{2}$$

where α_{ij} are $(1 \times K)$ vectors of parameters, and β_{ij} are scalars.

Examples of affine term structure models are the equilibrium single factor models of Vasicek (1977) and Cox, Ingersoll and Ross (1985). In these models the single factor Z_t is the instantaneous risk free rate, r_t. For $K > 1$, the models include the multifactor versions of the Vasicek and CIR models, and some generalisations. Further examples are discussed in Chen and Scott (1993) and Duffie and Kan (1996). Examples of the no-arbitrage class are the extended models of Vasicek and CIR considered by Hull

[4] Throughout the paper we will use the term "equilibrium" model for those models that endogenously determine the term structure by no-arbitrage arguments. Strictly speaking, not all affine models with constant parameters have a general equilibrium justification. Likewise, we will use the term "extended" for all models that start with an exogenously given yield curve. Sometimes these models are also called "evolutionary".

and White (1990). Campbell, Lo and MacKinlay (1997) develop the affine models in discrete time.

Define $Y_t(\tau)$ as the yield of a discount bond at time t with maturity τ. Frachot and Lesne (1993) show that with this specification of the factor dynamics no-arbitrage arguments imply a yield curve that is affine in the factors,

$$Y_t(\tau) = A_t(\tau) + B(\tau)Z_t \tag{3}$$

where $A_t(\tau)$ is a scalar and $B(\tau)$ a $(1 \times K)$ vector; both are functions of the structural parameters $\pi = (\alpha, \beta, \Phi_t, \Gamma)$ in the factor specification. In general the functions $A_t(\tau)$ and $B(\tau)$ are found by numerically solving a set of Ricatti differential (or difference in discrete time) equations. Next section considers the explicit functional relation for the special case of Gaussian factor models.

For pricing of interest rate derivatives the function $B(\tau)$ is the primary object of interest in term structure models. As shown in Heath, Jarrow and Morton (1992) and also explained in Hull (1996) all interest rate derivative securities can be priced with as inputs the "volatility" function $B(\tau)$, the factor covariance matrix Ψ_t, and the initial term structure. The purpose of the econometric analysis of yield curve models is the estimation of the function $B(\tau)$. For this the cross sectional model (3) suffices, and all we need is an assumption about the risk neutral factor dynamics.

The covariance matrix Ψ_t can often be estimated more precisely from time series data. Time series descriptions of the yield curve would focus on the factor process in (1). The general representation of the associated process under the actual probability measure reads

$$dZ_t = (\Phi_t - \lambda_t - \Gamma Z_t)\, dt + V(Z_t)d\tilde{W}_t \tag{4}$$

where λ_t is a vector of K prices of risk. Using time series data for statistical inference entails additional assumptions about the price of risk. In an equilibrium framework the risk prices are determined by preferences of agents, or more generally through a pricing kernel. While not every time variation in λ_t is allowed, the theory does not specify how they vary.

Before we develop an econometric model, it is useful to discuss the calibration practice of the extended models. The functions Φ_t and $A_t(\tau)$ have a time t subscript, but both are deterministic functions of time, whose time variation is fully determined by the observed term structure $Y_{t_0}(\tau)$ at date t_0, when the model was calibrated.

Over a short period the yield curve at time $t_0 + \Delta t$ will be close to the calibrated yield curve at time t_0 due to the smooth nature of the Brownian motion process. Over longer horizons the yield curve can wander away further from the initial yield curve. The variance of the distance is controlled by the integrated factor dynamics from t_0 to the current time t. Since the factor process is mean reverting the effect of the initial conditions will fade out gradually. Mean reversion implies that long term yields and forward rates both converge to a constant, and this in turn implies that, when t_0 is fixed, $\lim_{t\to\infty} \Phi_t = \tilde{\Phi}$ is a constant, and $A_t(\tau)$ will converge to a function of τ only. If the limit of Φ_t from the calibrated yield curve is the same constant as in the equilibrium formulation ($\tilde{\Phi} = \Phi$), then the yield curves generated by the extended models will converge to the yield curves from equilibrium formulations.

For the econometric inference we will use observed yield curves over a long period of time. If the extended model is calibrated only once, say in January 1960, then the yield curves in the nineties will only be affected by the initial calibration by the level parameter $\tilde{\Phi}$. In that case the extended model is in fact nothing but a restricted version of the equilibrium models. In the equilibrium models Φ is a free parameter, whereas it would be exogenously specified by more than thirty year old data in the extended model.

However, common practice in applications of the no-arbitrage models is repeated calibration, instead of a one time calibration. The model is calibrated every time we get a new term structure. With monthly data the functions Φ_t and $A_t(\tau)$ are updated every month, so that $t_0 = t - 1$ and $A_t(\tau)$ becomes a function of data from last month's term structure.

If, for the purpose of parameter estimation, the model is calibrated repeatedly, then the equilibrium and extended versions of the model only coincide if last month's term structure happens to be consistent with an equilibrium formulation. The extended model therefore allows for deviations from the unconditional equilibrium model that are expected to disappear gradually in a way that is consistent with the no-arbitrage condition. Because the intercept $A_t(\tau)$ depends on data from time $t - 1$, parameters are estimated from movements of the yield curve relative to its previous shape. In that sense it is a more flexible model for empirical work than the equilibrium model. When we develop the econometric model we will always consider the extended models being calibrated every month. In contrast to the equilibrium model, where the parameters are estimated from the shape of the yield curve, the extended models look at the cross section of changes in the term structure.

Using results in El Karoui, Frachot and Geman (1998) and Frachot and Lesne (1993) the affine model can be written in the representation

$$Y_t(\tau) = F_{t_0,t}(\tau) + a(t - t_0, \tau) + B(\tau)z_t, \tag{5}$$

where $F_{t_0,t}(\tau)$ is the forward rate at time t_0 for the contract period t to $t+\tau$, $a(t-t_0,\tau)$ is a function depending on the structural parameters, and z_t is a K-vector of factors (different from Z_t). Since for the econometric model we will assume $t_0 = t - 1$, the first argument of $a(t - t_0, \tau)$ does not depend on time, so that

$$Y_t(\tau) = F_{t-1,t}(\tau) + a(\tau) + B(\tau)z_t, \tag{6}$$

which is a representation with constant parameters. The extended model has the same factor loading $B(\tau)$ as the corresponding equilibrium model. But in the extended model the current yield curve $Y_t(\tau)$ is modelled conditional on the forward rate curve $F_{t-1,t}(\tau)$. The model has similar structure as the equilibrium model, but with yields in deviation of their corresponding forward rates instead of yield levels.

3 Gaussian Models

In general the transformation from factor dynamics to the affine yield curve parameters $A_t(\tau)$ and $B(\tau)$ is not available in closed form. The linear Gaussian case with uncorrelated factors is an example where closed form solutions are available, where comparison between equilibrium and extended models is straightforward, and where the structural parameters each have a clearly distinguishable function in the model. While the econometric model will be developed for the general affine class, the empirical analysis will concentrate on the Gaussian case with uncorrelated factors.

The single factor equilibrium Gaussian case is the Vasicek (1977) model. In this case $K = 1$ and $\alpha_{11} = 0$ in (1). The instantaneous spot rate r is the factor driving the yield curve, with risk adjusted diffusion process of the form

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \ , \tag{7}$$

where κ is the parameter of mean reversion, μ is the unconditional mean of r, and σ is the instantaneous standard deviation. A multiple factor generalization is given by

$$r_t = \sum_{j=1}^{K} z_{jt} \tag{8}$$

$$dz_{jt} = \kappa_j(\mu_j - z_{jt})dt + \sigma_j dW_{jt}, \tag{9}$$

where the Brownian motions W_{jt} are mutually uncorrelated. The associated yield curve model is represented by

$$Y_t(\tau) = A(\tau) + \sum_{j=1}^{K} b_j(\tau) Z_{jt}, \tag{10}$$

where

$$\begin{aligned} A(\tau) &= \sum_{j=1}^{K} \left((1 - b_j(\tau))\theta_j + \frac{(1-e^{-\kappa_j \tau})^2}{4\kappa_j^3 \tau} \sigma_j^2 \right) \\ b_j(\tau) &= \frac{1 - \exp(-\kappa_j \tau)}{\kappa_j \tau}, \end{aligned}$$

where $\theta_j = \mu_j - \frac{\sigma_j^2}{2\kappa_j^2}$. The shape of $B(\tau) = (b_1'(\tau), \ldots, b_K'(\tau))$ depends solely on the mean reversion parameters κ_j. Each element $b_j(\tau)$ is monotonically decreasing with $b_j(0) = 1$ and $b_j(\infty) = 0$. As in all affine models, the long term discount yield $Y(\infty)$ is a constant. At every date the yield curve must converge to the same constant, irrespective of the initial location and shape. Note that in the multifactor model not all parameters are identified. Reparameterising by $Z_{jt} = z_{jt} - \theta_j$, one can verify that we can only identify $\theta = \sum_{j=1}^{K} \theta_j = \lim_{\tau \to \infty} Y_t(\tau)$, but not the individual θ_j.[5] Also note that the model is not only affine in the factors Z_{jt}, but also in the parameters θ and σ_j^2. For cross sectional estimation, the only nonlinearity is in κ_j.

The parameters θ and σ_j^2 only show up in $A(\tau)$, so that $B(\tau)$ solely depends on the mean reversion. Since the factor loadings (or volatility structure) $B(\tau)$ is the object of interest for the purpose of derivative pricing, our main concern will be the estimation of the mean reversion parameters. This is a specific feature of the Gaussian models. In the general affine model, and also in the CIR model, the factor loadings depend on all structural parameters.

The extended Vasicek is similar to (7), but now μ is replaced by μ_t. The parameters μ_t fluctuate in such a way that they are consistent with the observed term structure characteristics at some initial date t_0. Calibrating the term structure at some initial time t_0, Hull and White (1993) derive expressions for the coefficients μ_t and $A_t(\tau)$ for $t > t_0$. Letting $t_0 = t - 1$, and defining h as the length of the interval between successive observations we use their results to obtain an expression analogous to (6),

$$Y_t(\tau) - F_{t-1,t}(\tau) = a(\tau) + B(\tau)(r_t - f_{t-1,t}), \tag{11}$$

[5] See Dai and Singleton (1997) for a detailed analysis of identification and alternative parameterizations.

where

$$a(\tau) = (1 - e^{-2\kappa h})\frac{(1 - e^{-\kappa\tau})^2}{4\kappa^3\tau}\sigma^2 , \qquad (12)$$

$F_{t-1,t}(\tau)$ is the forward rate defined below (5), and $f_{t-1,t}$ is the instantaneous forward rate at time $t-1$ relating to time t. The variables $Y_t(\tau) - F_{t-1,t}(\tau)$ and $r_t - f_{t-1,t}$ can be interpreted as unexpected shocks under the risk neutral probability measure. The main differences between the standard and the extended Vasicek models are the restrictions on the levels of the yields, which appear in the standard Vasicek model through the parameter θ, but not in the extended version.

From the results of Frachot and Lesne (1993) we obtain a multifactor generalisation of (11) as

$$Y_t(\tau) - F_{t-1,t}(\tau) = a(\tau) + \sum_{j=1}^{K} b_j(\tau) Z_{jt} , \qquad (13)$$

where $b_j(\tau)$ is the same as in (10), and

$$a(\tau) = \sum_{j=1}^{K}(1 - e^{-2\kappa_j h})\frac{(1 - e^{-\kappa_j\tau})^2}{4\kappa_j^3\tau}\sigma_j^2 .$$

So far we only discussed the risk neutral factor process. To include time series properties we could make the additional assumption that the price of risk λ is a constant. For example, internal coherence of the one factor model then requires that the time series process of r_t must be similar to the spot rate process, given in (7),

$$dr = \tilde{\kappa}(\tilde{\mu} - r)dt + \tilde{\sigma} d\tilde{W}_t \qquad (14)$$

The tildes on top of the structural parameters indicate that this process is specified under the original probability measure with Brownian motion $\tilde{W}$ instead of the risk neutral processes in (7). If the model is correctly specified the structural parameters differ from the risk neutral parameters only by the parameter $\lambda = \mu - \tilde{\mu}$, while $\tilde{\kappa} = \kappa$ and $\tilde{\sigma}^2 = \sigma^2$. The equality of the "implied" and "actual" parameters are testable restrictions.

For the no-arbitrage models the time series process of r_t itself has time varying parameters. But, assuming a constant price of risk, equation (11) implies that the deviation between the short rate and the lagged forward rate $r_t - f_{t-1,t}$ must be serially uncorrelated under the actual factor dynamics. The mean of $r_t - f_{t-1,t}$ under the actual probability measure is not determined by the model, and depends on the price of risk λ. As it is the innovation in the instantaneous spot rate, the variance must be equal to the conditional variance of r_t over a horizon of length h,

$$\mathrm{Var}(r_t - f_{t-1,t}) = \tilde{\sigma}^2 \left(\frac{1 - \exp(-2\tilde{\kappa}h)}{2\tilde{\kappa}} \right) \tag{15}$$

As in the equilibrium version of the model, consistency between the time series and cross sectional dimension implies $\tilde{\kappa} = \kappa$ and $\tilde{\sigma}^2 = \sigma^2$. However, since the variance of $r_t - f_{t-1,t}$ is the only moment that depends on these two parameters, they cannot be separately identified.

4 Why Pooling?

Before developing a fully articulated econometric model, we briefly illustrate the problems in cross sectional estimation, and give the intuition why a panel data analysis will be helpful. For the example we consider the single factor Vasicek model.

Suppose the Vasicek model is true with parameters $\kappa = 0.04$, $\sigma = 2.5\%$ and $\theta = 0$. Figure 1 shows three yield curves starting with spot rates at 4%, 8% and 12% respectively. The yield curves are drawn for maturities between 0 and 10 years. For the selected parameter configuration all three yield curves are upward sloping (only the curve starting at $r = 12\%$ has a slight hump) and almost parallel, although each of them must be downward sloping eventually, and converge to $Y_t(\infty) = \theta = 0$. Even though the Vasicek model implies that all yield curves converge to the same long term yield, this convergence is not visible over horizons up to 10 years when the mean reversion parameter κ is very small. A value of κ so close to the unit root is also plausible from a time series perspective (see for example Chan, Karolyi, Longstaff and Sanders (1992)).

In cross sectional estimation the parameters for the yield curve at time t are r_t, κ, σ, and θ. When estimated period by period, they can be different for each of the three curves. If the data are pooled, only r_t is different for the three curves, while κ, σ and θ are constant. In the example the yield curve data are without any measurement error, so that we can of course uniquely recover the true parameters of the data generating process. To show the problems with cross sectional estimators we simulate data by adding a small amount of measurement noise to the data, in the order of 8 basispoints at the short end to 3 basispoints at the long end. Errors with such a small standard deviation are not visible to the eye on the scale of figure 1.

In our Monte Carlo experiment parameters are estimated by nonlinear ordinary least squares. In effect, the Vasicek model is only nonlinear in κ. Conditional on κ, the estimator reduces to ordinary least squares with fixed regressors. The results in table 1 are based on 1500 replications. We find that the expected value of the least squares estimate $\hat{\kappa}$ increases from 0.066 (when $r_t = 4\%$) to 0.201 (when $r_t = 12\%$). The cross sectional estimator is biased towards too much mean reversion, and the bias increases with the level of interest rates. Since the least squares estimator is consistent, the bias will disappear, but only when we add longer maturities beyond the ten year horizon.

The bias in $\hat{\kappa}$ directly affects the estimates of volatility. The estimates of σ^2 are biased downwards, and show enormous variation. The parameter estimates are as erratic as reported in the empirical studies referred to in footnote 1. As simple as it is, the Vasicek model is already overparameterised in cross sectional estimation. The cross sectional estimates can have a tendency to overestimate the mean reversion.

Next we pool the three cross sections by simply stacking the data for the three yield curves. The six parameters (r_1, r_2, r_3, κ, σ^2, and θ) are estimated jointly by least squares. Pooling the three cross sections is extremely effective. Because the three yield curves are at very different levels, the parameters are estimated without any problem. There is no bias in $\hat{\kappa}$, and its standard error is negligible (reduced by a factor of 20 compared to the first cross section, and even a factor of 100 compared to the third cross section). With κ estimated almost without error, it is even possible to estimate the variance σ^2 from the shape of the yield curve. The long run rate θ is still estimated imprecisely, but this is because it is almost unidentified when κ is small (see figure 1), and because we do not use maturities longer than 10 years.

The lower part of table 1 presents simulation results when the amount of noise is increased substantially. More noise leaves the estimator of $\hat{\kappa}$ unbiased, although its standard error of course gets larger. With more noise, estimation of θ becomes hopeless. However, θ is not a parameter of interest in applications of the Vasicek model.

5 Econometric Model

The ordinary least squares estimator in the example in section 4 above is not necessarily the most efficient estimator, since it ignores the measurement error structure. The error covariance properties of the empirical data will also affect the statistical inference. We therefore now turn to an econometric model of the error term.

The panel structure in the general affine model (3) is apparent, since the model has a time series dimension t and a cross sectional dimension τ. An important difference with most panel data models discussed in the econometric literature is the natural ordering of observations in the cross sectional dimension.[6] The model is continuous in both dimensions. However, data are sampled in discrete time, and for a discrete number of maturities. In a time series analysis the typical data sets are daily, weekly or monthly and can cover a period of more than 30 years. The number of cross sectional observations is limited by the number of traded bonds. If the data consist of discount bond prices, these are usually obtained by spline methods.[7] In that case the "observed" — in fact constructed — term structure will be a smooth curve in the maturity dimension. It is this type of data that we have in mind when we develop the econometric model. We consider a dataset of discount yields observed at dates $t = 1, \ldots, T$ and with maturities $\tau_1, \ldots, \tau_N$. If the yield curve is constructed using spline methods, it is of course possible to select as many points on the yield curve as we wish, but since neighboring yields will be almost perfectly correlated, the information contents will reach an upper bound determined by the number of parameters in the spline function.

5.1 Specification

All term structure models imply that there is a deterministic relation between several yields. A perfect fit with N maturities is achieved by the inclusion of N factors. When the number of factors is less than the number of maturities, the model implies $N - K$ deterministic linear relations between the yields. This holds both for the equilibrium as well as the extended models. Although the extended models provide a perfect fit for all N points on the yield curve at some time t_0, for $t > t_0$ there are only K state variables driving all N yields.

To avoid such singularities many studies add statistical noise to equilibrium formulations of the yield curve.[8] The econometric model of the yield curve model then

[6] See Baltagi (1995) and Hsiao (1986) for textbook treatments of panel data models.

[7] See McCulloch (1975) and Nelson and Siegel (1987).

[8] See for example Brown and Dybvig (1986), Gibbons and Ramaswamy (1993), Frachot, Lesne and Renault (1995), DeJong (1997), Lund (1997) and others for noise in an equilibrium model. Bliss and Ritchken (1996) is an example of an error term in an extended model. Jacquier and Jarrow (1998) provide an extensive motivation why one should add noise in a contingent claims model, and how to interpret the error term.

becomes

$$Y_t = A_t + BZ_t + e_t \tag{16}$$

where Y_t, A_t and e_t are vectors of length N, and B is an $(N \times K)$ matrix. The i^{th} element in Y_t corresponds to maturity τ_i, and maturities are ordered $\tau_1 < \ldots < \tau_N$. The error term e_t captures all forms of measurement and specification error, but its properties are not part of the economic theory.[9] If the model is to be of any practical use, the error variance should be small relative to both the cross sectional and the time series variation in yields. Otherwise, since the error deals with unmodelled phenomena, there are no apparent restrictions on its covariance structure.

The term structure data are observed as a smooth function of maturity, to which we fit a curve that is also a smooth function of τ, as illustrated in figure 2. Therefore we assume that the correlation between error terms $e(\tau_i)$ and $e(\tau_j)$ depends on the distance between the terms to maturity:

$$\text{corr}(e_t(\tau_i), e_t(\tau_j)) = \phi^{|\tau_i - \tau_j|}, \tag{17}$$

with $0 < \phi < 1$. Yields that are very close show high correlation, whereas yields that are far apart are less correlated. The specification resembles a cross sectional AR(1) error term.

Besides cross sectional correlation we also account for possible cross sectional heteroskedasticity of the error terms. The theoretical model can be written either in terms of log-prices or in yields. A transformation from yields to log prices $\ln P_t(\tau) = -\tau Y_t(\tau)$ induces maturity specific variances with a proportionality factor equal to τ^2. For that reason we specify the variance of the error term $e_t(\tau)$ as a function of τ:

$$\text{Var}(e_t(\tau)) = \omega^2 \tau^{-2d}, \tag{18}$$

where ω^2 is a scale parameter, and d determines the sensitivity of the variance for the term to maturity. If $d = 0$ then the error terms are homoskedastic in a model for the yields. In case $d = 1$, the model is homoskedastic in a regression model for (log-) bond prices. The heteroskedasticity implies a weighting scheme on the maturities: with $d = 0$ all yields have equal weights; with $d > 0$ more emphasis is put on long term yields. We estimate d along with the other parameters.

[9] See Renault (1996) for examples of an endogenous error term in the context of an option pricing model. In order to obtain a nondegenerate error term one must start at the theoretical level with at least as many state variables as there are observations in the cross section. It is unlikely however, that one would ever want to go beyond three or four factors. Part of the statistical specification of the model will always remain somewhat adhoc.

In matrix form the covariance structure for the cross sectional error terms is

$$\Sigma = \omega^2 S(\phi, d) \tag{19}$$

where S is a matrix of order $(N \times N)$ with typical element $s_{ij} = (\tau_i \tau_j)^{-d} \phi^{|\tau_i - \tau_j|}$.

As a further extension of the statistical model of the error term we allow for first order autocorrelation in the error terms of (16) like in Chen and Scott (1993),

$$e_t(\tau) = c e_{t-1}(\tau) + \epsilon_t(\tau) , \tag{20}$$

where $\epsilon_t(\tau)$ is uncorrelated over time, and has the covariance structure specified in (19). We restrict the autocorrelation parameter c to be the same for each maturity τ in order to preserve the smoothness of the cross sectional error process.

The possible autocorrelation in the errors makes the equilibrium model flexible enough to approximate the extended model. If in the equilibrium model the autocorrelation parameter c approaches one, we obtain a model in first differences. In that case $B(\tau)$ is estimated from yield changes, while the level function $A(\tau)$ becomes unidentified. But empirically long maturity forward rates $F_{t-1,t}(\tau)$ are almost identical to the corresponding spot rates $Y_{t-1}(\tau)$. Yield changes $\Delta Y_t(\tau)$ will thus be very similar to the forward deviations $Y_t(\tau) - F_{t-1,t}(\tau)$ that are input to the extended models. Since $B(\tau)$ is identical in the equilibrium and extended models, an equilibrium model in first differences ($c = 1$) will be very similar to an extended model with $c = 0$, except at the short maturities.

Altogether the error specification consists of the parameters $\zeta = (\omega, \phi, d, c)$. The specification of the error term differs from other models in the literature in several respects. First, the error term has been added at the level of the discount price function, and not on the original traded individual (coupon) bond prices as in Brown and Dybvig (1986), Schotman (1996) and DeMunnik and Schotman (1994). Adding the error term at the level of discount bond prices has been proposed by Gouriéroux and Scaillet (1994), who motivate this choice from no-arbitrage conditions and invariance properties in modelling portfolios of bonds. However, Gouriéroux and Scaillet (1994) assume that the errors are uncorrelated across maturities, whereas our specification explicitly takes into account the smooth nature of observed yield curves.

Chen and Scott (1993) first eliminate the unobserved factors Z_t from (16), by using the exact relation between the factors and K different yields with maturities τ_{0j}, and then add noise to the equations for all other maturities. This would imply that the particular yields with maturity τ_{0j} are always fitted exactly, *i.e.* $e_t(\tau_{0j}) = 0$ for $j = 1, \ldots, K$. Apart from this singularity, Chen and Scott (1993) allow for a fully unrestricted covariance matrix Σ. Bliss and Ritchken (1996) introduce the same singularity in their analysis of the extended Vasicek model, and also tightly parameterise the error covariance matrix. Frachot, Lesne and Renault (1995) and Bliss and Ritchken (1996) note that the choice of the pivotal maturities τ_{0j} is very influential for the empirical results. Therefore De Jong (1997) and Frachot, Lesne and Renault (1995) eliminate the unobserved state variables in a way that is invariant with respect to the maturity τ_{0j}. The number of unknown parameters in Σ in their specifications is of order N^2. Such a general unrestricted specification limits the number of maturities that can be included in the empirical analysis. In most of the empirical studies N is therefore only 4 or 5. Our tight parameterization of Σ allows N to be large, which helps to increase the power of the cross sectional tests. Eventually the strong positive correlation between neighboring maturities will put an upper bound on the information contents in the data.

Although tightly parameterised, our error covariance matrix is less restrictive than in other papers that use a broad spectrum of maturities. Gibbons and Ramaswamy (1993) assume absence of serial correlation in the measurement error. Many other studies (for example Geyer and Pichler (1997), Lund (1997)) assume absence of cross sectional correlation in the errors.

5.2 Estimation and Testing

The general model for both equilibrium and extended versions of the affine class is (16). For the estimation of the model standard panel data methods are applicable. The least restrictive assumption on the K factors Z_t is to treat them as a time series of unknown parameters, *i.e.* as fixed effects. The fixed effects approach is purely cross sectional; no time series information is used for the estimation of the structural parameters π of the model, and we therefore do not need assumptions on the dynamics of Z_t. Treating Z_t as parameters, the model in (16) can already be estimated on data for a single point in time t as in section 4. The panel data aspects arise from pooling the data for several cross sections.

We use Quasi Maximum Likelihood to estimate the structural parameters π, the error parameters ζ and the fixed effects $\{Z_t\}_{t=1}^{T}$. To handle the autocorrelation in the error terms we adopt the following transformation for any time series variable X_t

$$\begin{aligned} X_t^* &= (1-c^2)^{\frac{1}{2}} X_1 && t=1 \\ &= X_t - cX_{t-1} && t=2,\ldots,T \end{aligned} \tag{21}$$

Using (21) on yields, factors, $A_t(\tau)$ and the errors we transform (20) to the model

$$Y_t^* = A_t^* + BZ_t^* + \epsilon_t, \tag{22}$$

for which we have the quasi-loglikelihood function

$$\begin{aligned} \ln L(\pi,\zeta,\mathbf{Z}) &= -\tfrac{1}{2}NT\ln(\omega^2) - \tfrac{1}{2}T\ln|S| + \tfrac{1}{2}N\ln(1-c^2) + \\ &\quad -\frac{1}{2\omega^2}\sum_{t=1}^{T}(Y_t^* - A_t^* - BZ_t^*)'S^{-1}(Y_t^* - A_t^* - BZ_t^*), \end{aligned} \tag{23}$$

where $\boldsymbol{Z} = (Z_1,\ldots,Z_T)$. The likelihood can be concentrated analytically with respect to the fixed effects $\boldsymbol{Z}$ and the scale parameter ω^2. This is the crucial element of an affine model. For models not in the affine class the fixed effects cannot be concentrated out analytically, which would add T parameters to the non-linear numerical optimization of the likelihood function in (23), rendering these models intractable for analysis by conventional methods. For ω^2 the maximum likelihood estimator is given by

$$\hat{\omega}^2 = \frac{1}{NT}\sum_{t=1}^{T}(Y_t^* - A_t^*)'M(Y_t^* - A_t^*)\,, \tag{24}$$

where $M = S^{-1} - S^{-1}B(B'S^{-1}B)^{-1}B'S^{-1}$. Substitution of (24) in (23) gives the concentrated likelihood function

$$\ln L(\pi,c,d,\phi) = -\tfrac{1}{2}NT\ln\hat{\omega}^2 - \tfrac{1}{2}T\ln|S| + \tfrac{1}{2}N\ln(1-c^2) \tag{25}$$

which depends on only a small number of parameters, and can be easily maximized by numerical optimization routines.

The structural parameters only affect the first term in (25). It is instructive to consider this term in more detail. For the equilibrium models A_t does not depend on t. Let $\boldsymbol{V}$ be the sample covariance matrix of Y_t^* (which depends on c), and let $\bar{Y}$ be the sample mean of Y_t. Equation (24) can be rewritten as

$$\hat{\omega}^2 = \frac{1}{N}\left(\mathrm{tr}(\boldsymbol{MV}) + (\bar{Y}-A)'\boldsymbol{M}(\bar{Y}-A)\right)\,, \tag{26}$$

ignoring a term of order $(Y_T - Y_1)/T$. For the Gaussian model this decomposition implies that the function $A(\tau)$, and therefore the parameters θ and σ_j^2 are determined by the unconditional sample means $\bar{Y}$, *i.e.* the average shape of the yield curve. The function $B(\tau)$ — and thus mean reversion parameters κ_j — are identified through all first and second moments of the transformed yields. QML defines a weighting scheme based on all possible $\frac{1}{2}N(N+3)$ moment conditions.

A similar decomposition for the extended Vasicek model in (11) shows that the volatility parameter is only determined through the convexity in $\bar{Y}-\bar{F}$ as a function of τ ($\bar{F}$ is the sample average of the N-vector with elements $F_t(\tau_i)$). Without using time series information all the information on σ^2 has to be extracted from the sample means of the deviation between the yield and the forward rate. Since this is not likely to be very informative data, estimates of σ^2 are presumably imprecise in the extended models.

The cross sectional specification of any affine term structure model is testable by comparing $A_t(\tau)$ and $B(\tau)$ with a less restrictive specification. To test for deviations from the theoretical model we augment the functional forms $A_t(\tau)$ and $B(\tau)$,

$$
\begin{aligned}
\tilde{A}_t(\tau) &= A_t(\tau) + \gamma' g(\tau) \\
\tilde{B}(\tau) &= B(\tau) + Dg(\tau)\,,
\end{aligned}
\tag{27}
$$

where $g(\tau)$ is an L-vector of functions of τ, and γ and D are an L-vector and $(K \times L)$ matrix of parameters.

This tests considers the specification of the model keeping the number of factors constant. The number of factors itself can be examined by comparing with a model that has more factors. A formal test for the number of factors is not trivial due to the increasing number of incidental parameters in $\boldsymbol{Z}$. Every additional factor k introduces T new parameters in $\{Z_{kt}\}$, invalidating standard large T asymptotic inference. When the number of factors K equals the number of maturities N, the model will fit perfectly, and the likelihood function will go to infinity.

Unlike in standard factor models, or in principal components analysis, the "specific" risk terms $e_t(\tau)$ in (16) are not necessarily mutually uncorrelated. In fact, as discussed in section 2, they most likely are not, because of the smooth nature of the term structure data. Finding more than one factor in the covariance structure of interest rate data does not necessarily invalidate single factor term structure models. It is the shape of the factor loadings $B(\tau)$ that constitute the testable implications of the model.

6 Data

The data set consists of a panel of discount yields, which are constructed using the McCulloch (1975) spline procedure from US government bond data, available from the CRSP tapes. The data are monthly observations that span the period from January 1970 until December 1994. At every time t the yield curve is represented by a cross section of $N = 16$ yields. The maturities are one through six months and one year until ten years.

Figure 3 shows the full data panel. Most striking is the predominance of parallel shifts of the yield curve. Apart from noisy behaviour at the short end, yield changes are almost horizontal. Yields with different maturities are therefore heavily correlated, both in levels and in deviation from lagged forward rates.

Summaries are presented in figure 4. The average yield curve in figure 4A is concave, as it should be in the Vasicek model in order to obtain a positive estimate of volatility. In deviation of the forward rate the concavity does not hold at the very short maturities of one- and two months in figure 4B.

Both for levels and changes the volatility in figures 4C and 4D decreases with maturity, as they should when interest rates are mean reverting. But the term structure of volatilities for the levels is flatter than that for the changes. This is primarily due to the initial steep decrease in volatility for the very short term rates (maturities six months and less). Very short rates are different from the rest of the term structure. Volatility at longer maturities is still far from zero, indicating that the ten year interest rate is not a proxy for the infinite horizon yield that must be constant over time in the affine model. The observed yield curves have not converged to a single yield for long maturities.

The term structure of volatilities in figure 4D for first differences $Y_t(\tau) - Y_{t-1}(\tau)$ is almost identical to the volatility structure for the deviations between yields $Y_t(\tau)$ and lagged forward rates $F_{t-1,t}(\tau)$. This implies that the extended model is in effect a model of yield changes. This is brought out more clearly in the scatter diagram in figure 6.

Figure 5 shows yield curves ordered by the level of the ten year rate. For a one factor model, all curves with a common ten year rate should be identical apart from measurement error. Instead the curves spread out at the short maturities. One interpretation of these data is a second factor with strong mean reversion that only has an effect on short term yields. Another possibility is that measurement error for short maturities is relatively high.

7 Empirical Results

In the empirical analysis we consider four different models: equilibrium and extended Gaussian models, with either one or two factors. Parameter estimates are reported in the first column of table 2. We first discuss the Vasicek single factor equilibrium model in detail, and then proceed with the other three models.

The parameter of interest in the Vasicek model is the implicit mean reversion κ, since it completely determines the volatility function $B(\tau)$. The panel model pools the yield curves for many different months. During the sample period short term interest rates have fluctuated between 5 and 15 percent, *i.e.*, the yield curves start from very different levels of the instantaneous spot rate. When all these yield curves are forced to converge to a single infinite yield, the only way that any single factor model can achieve this, is by setting the mean reversion parameter at a very small value. The point estimate implies a monthly autoregressive coefficient of 0.997, or equivalently a halflife of shocks equal to $\ln 2/\kappa = 17$ years. From the cross sectional perspective interest rates have to be near integrated series. Although the mean reversion is low, it is estimated very precisely and is significantly different from zero, with a t-statistic that is larger than 8.

At the very long end of the yield curve, the estimate of the infinite yield θ is negative and imprecise. It appears that this parameter is not identified in the data. The poor estimates are related to the low mean reversion, since the constant infinite maturity yield becomes unidentified when κ is close to zero. We simply have no reliable data about very long term interest rates.[10,11,12] Although the infinite yield is negative, the model generates upward sloping, almost parallel, yield curves for maturities up to 10 years. Due to the low mean reversion the term structures are far

[10] The CRSP data set does contain bonds with a maturity of 30 years. We included these bonds, when we estimated spline functions to create discount yields. However, the spline function shows very large standard errors at maturities beyond ten years, indicating that the yield curve data are measured with much error for these long maturities. For that reason we did not extend our discount yield data beyond the ten year maturity.

[11] The CIR model yields similar results. Although mean reversion is even lower at 1.25e-6 and the infinite yield is positive at 6.35 by construction, the fit of the term structure is almost identical.

[12] Setting $\theta = 0$ instead of the ridiculously large negative value makes absolutely no difference for the estimate of κ. It only has some effects on σ.

from convergence to their common infinite horizon yield at a maturity of ten years. It might be considered a serious drawback of the Vasicek model, like all other one factor models, that it implies one constant infinite yield. But even in a multifactor world, it is not trivial to construct an equilibrium model without this feature. Variability of long maturity interest rates will always imply low mean reversion.[13]

Other panel data studies, referred to in the introduction, also incorporate time series information. Instead of estimating the factor as an unknown parameter, they treat it as a latent variable that follows the Gaussian process (14). The parameters κ and σ appear in both the cross sectional as well as the time series dimension of the model, in other words in both the risk neutral and actual dynamics. That way these studies use all possible information in the data at the cost of making the assumption of a constant price of risk. But the additional information in the time series seems weak. When we use the Kalman filter to estimate the model by maximum likelihood with r_t as a latent variable, the parameter estimates remain unchanged. Our point estimates are $\hat{\kappa} = 0.039$ and $\hat{\sigma} = 0.337$ with standard errors that are only slightly less than those in table 2. The cross sectional information dominates the time series information in the data. This should not come as a surprise, since time series data are known not to be informative on mean reversion. That was one of the conclusions of Chan, Karolyi, Longstaff and Sanders (1992). When time series and cross sectional parameters are allowed to differ in the way discussed below (14), we indeed find that the time series mean reversion $\tilde{\kappa} = 0.21$ gets a standard error of 0.12, and is not significantly different from zero.

The diagnostic tests give many indications that the one factor moded is misspecified. First, the likelihood ratio test rejects the structure of $A(\tau)$ and $B(\tau)$ imposed by the Vasicek model. The deviations are illustrated in figures 7A and 8A. Figure 7 shows the average observed term structure and the average of the estimated term structure, given by $A(\tau) + B(\tau)\bar{z}$, where $\bar{z}$ is the sample mean of the estimated factor. On average the actual term structure is steeper than what would be consistent with the one factor Vasicek model. Figure 8 shows the volatility structure. The estimated volatility is computed as the standard deviation of $B(\tau)z_t$. Short term rates are much more volatile than implied by the model. Another way to interpret figure 8 is as a measure of fit. The ratio of fitted to actual variance, $\text{Var}(B(\tau)z_t)/\text{Var}(Y_t(\tau))$ for different maturities can be used as an R^2 measure of goodness of fit. The explained

[13] See for example the theoretical discussion in El Karoui, Frachot and Geman (1998).

variance quickly rises from 31% for the one month interest rate to over 99% for yields with maturities 3 years and longer.

Second, the estimate of the cross sectional heteroskedasticity parameter d implies that parameter estimates are mostly determined by data from the long end of the maturity spectrum, and that the model does not fit the short end at all. With $\hat{d} = 0.75$, the relative weight of the ten year yield compared to the one month rate is $(\frac{10}{1/12})^{\hat{d}} \approx 36$, against equal weighting when $d = 0$.

Third, the estimate of ϕ states that cross sectional dependence of the errors is strong. For yields that differ only one month in time to maturity, the correlation coefficients are about $\phi^{1/12} = 0.97$. There is a strong common component left in the errors. This common component also has strong autocorrelation, but is far away from the unit root.

Fourth, figure 9 depicts the implied instantaneous spot rate, *i.e.*, the estimated fixed effect $\hat{r}_t$, together with the observed one month rate (as a proxy for the short rate). The implied rate is less volatile than the observed rate. It looks as if the actual one month interest rate contains a lot of transitory noise. The same transitory noise is also visible in the yield curve changes in figure 3.

All these four empirical diagnostics point at a second factor, which has stronger mean reversion than the dominant first factor, and which therefore will mostly affect the shorter maturities. The second column in table 2 show the estimation results. As anticipated we find one factor with very low mean reversion, like in the one factor case. The second factor is much more volatile but strongly mean reverting. This model fits the data much better. The restrictions on $A(\tau)$ and $B(\tau)$ can not be rejected, and therefore the actual and fitted first and second moments match almost perfectly in figures 7 and 8. The two factors account for more than 95% of all variation in interest rates. The remaining error terms are small. Moreover, most of the cross sectional correlation in the errors has disappeared.

The high mean reversion of the second factor is entirely consistent with the time series behavior of short term interest rates. Both Andersen and Lund (1996) and Balduzzi, Das and Foresi (1998) find that the short term interest rate can be described as a time series that quickly reverts towards a slowly changing mean. Dai and Singleton (1997) show that such a model, which they call a "cascade" model, is just an alternative parameterisation of the two-factor model.

A disturbing diagnostic is that the residual autocorrelation remains. The second factor only fixes the misspecification at the short end, but does not affect the longer term yields. Since most of the autocorrelation comes from the longer term yields, the second factor can not take that away.[14] When the theoretical yield curve does not match the observed yield curve, the error persists for some time. That is exactly the type of behavior that can be modelled by an extended model. An extended model starts from an exogenously given yield curve, and assumes that movements away from the initial conditions satisfy a no-arbitrage constraint. If the current yield curve is not consistent with an equilibrium model, then the error will only partly be corrected in next period's yield curve. When an equilibrium is fitted to data from an extended model with an arbitrary initial condition, the error terms will exhibit strong positive autocorrelation. Estimation of an extended model is therefore directly motivated by the results from the equilibrium model.

The estimate of the mean reversion in the extended Vasicek model in table 2 is close to the estimate for the equilibrium version of the model. Like the equilibrium model, the one factor extended Vasicek model has problems fitting the short term interest rate movements. The observed volatility is much higher than the volatility $\text{Var}(B(\tau)(r_t - f_{t-1,t}))^{\frac{1}{2}}$ implied by the estimated factor.

In the two factor extended model, the first factor is almost a random walk. Since the equilibrium and extended Vasicek models are not nested, they cannot be formally tested against each other. The two models aim at fitting different sets of moments in the data. Still the likelihood value of the extended two-factor model falls short of the likelihood value of the equilibrium model. Other diagnostics also suggest that the equilibrium version of the model fits the data better.

A major difference between the extended and equilibrium models is the autocorrelation in the residuals. First, for the equilibrium model, specified in levels, the estimated autocorrelation is about $c \approx 0.75$. For the extended Vasicek model, which is practically equivalent to taking first differences of the yields, the autocorrelation coefficient becomes negative at about $c \approx -0.30$. Negative autocorrelation is an indication that working with monthly changes in yields leads to 'overdifferencing'. The extended Vasicek model calibrates too often! What is missing in the extended Vasicek model is an "error correction" factor that measures the deviation of the current term structure from the expected yield curve based on an equilibrium model. The levels

[14] This of course means that our parameterisation of the measurement error structure is too restrictive, since it imposes the same autocorrelation for all maturities. But this does not affect the consistency of the parameter estimator, and only means that a more efficient estimator exists.

of interest rates contain useful information about the parameters that is ignored in the extended model. This could explain why parameter estimates are more precise for the equilibrium version of the model.

Second, the estimated volatility structure in figure 8 indicates that the two-factor model still does not account for the relatively high volatility at the short end of the term structure. This is directly related to the low mean reversion in the second factor. The low value for $\hat{\kappa}_2$ arises from the covariance structure of $Y_t(\tau) - F_{t-1,t}(\tau)$. The correlations between $Y_t(\tau) - F_{t-1,t}(\tau)$ are generally lower than between the yield levels $Y_t(\tau)$, even after correcting for autocorrelation as in $Y_t^*(\tau)$. The data need the second factor to improve the fit of movements in long term rates.

8 Conclusions

We have proposed a fixed effects panel data model for the term structure of interest rates. This framework allows us to estimate the risk neutral process of the factors without making additional assumptions on the actual time series behavior of the factors. In this framework we treat single and multifactor models as well as equilibrium and extended term structure models. Two issues are important in the econometric model. First, we consider the affine class of term structure models, since this class allows the use of linear panel data methods. Second, the model takes into account the natural ordering of interest rates in the maturity dimension to parameterise the covariance matrix of the error terms.

Pooling results in very sharp point estimates of the structural parameters. Most salient is the implied mean reversion in a one factor Gaussian model, which is very low but statistically different from zero. Such precise estimates are not available from time series analysis of 25 years of monthly data. Moreover the equilibrium version model fits very tight for maturities larger than one year. At the very short side of the yield curve, we find a lot of transitory noise, which requires a second factor. For the extended Gaussian models we find strong negative autocorrelation as evidence of overdifferencing. The extended models ignore useful information in the levels of the yield curve.

References

ADDOLORATO, F. AND A. BERARDI (1994), Long-Term Yield Volatility and the Term Structure: Evidence for Italian Treasury Bonds, Working Paper, London Business School.

ANDERSEN, T.G. AND J. LUND (1996), The Short Rate Diffusion Revisited: An Investigation guided by the Efficient Method of Moments, Working paper (Aarhus School of Business).

BABBS, S.H. AND K.B. NOWMAN (1998), Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis*, Forthcoming.

BALDUZZI, P., S.R. DAS, AND S. FORESI (1998), The Central Tendency: A Second Factor in Bond Yields, *Review of Economics and Statistics*, **80**, 62–72.

BALTAGI, B.H. (1995), *Econometric Analysis of Panel Data*, Wiley, Chicester.

BLISS, R.R. AND P. RITCHKEN (1996), Empirical Tests of Two State-Variable Heath-Jarrow-Morton Models, *Journal of Money Credit and Banking*, **28**, 452–476.

BROWN, R.H. AND S.M. SCHAEFER (1994), The Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross Model, *Journal of Financial Economics*, **35**, 3–42.

BROWN, S.J. AND P.H. DYBVIG (1986), The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance*, **41**, 617–630.

CAMPBELL, J.Y., A.W. LO, AND A.C. MACKINLAY (1997), *The Econometrics of Financial Markets*, Princeton University Press.

CHAN, K.C., G.A. KAROLYI, F.A. LONGSTAFF, AND A.B. SANDERS (1992), An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, **47**, 1209–1227.

CHEN, R.R AND L. SCOTT (1993), Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income*, **3**, 14–31.

COX, J.C., J.E. INGERSOLL, AND S.A. ROSS (1985), A Theory of the Term Structure of Interest Rates, *Econometrica*, **53**, 385–407.

DAHLQUIST, M. AND L.E.O. SVENSSON (1994), Estimating the Term Structure of Interest Rates with Simple and Complex Functional Forms: Nelson & Siegel versus Longstaff & Schwartz, Seminar paper 565, Institute of International Economic Studies, University of Stockholm.

DE JONG, F. (1997), Time-Series and Cross Section Information in Term Structure Models, Working Paper, Tilburg University.

DE MUNNIK, J.F.J. AND P.C. SCHOTMAN (1994), Cross Sectional Versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market, *Journal of Banking and Finance*, **18**, 997–1025.

DUAN, J.C. AND J.G. SIMONATO (1995), Estimating Exponential-Affine Term Structure Models by Kalman Filter, Working paper, McGill University.

DUFFEE, G.R. (1996), Idiosyncratic Variation in Treasury Bill Yields, *Journal of Finance*, **51**, 527–551.

DUFFIE, D. AND R. KAN (1996), A Yield Factor Model of Interest Rates, *Mathematical Finance*, **6**, 379–406.

EL KAROUI, N., A. FRACHOT, AND H. GEMAN (1998), On the Behavior of Long Zero Coupon Rates in a No Arbitrage Framework, *Review of Derivatives Research*, **1**, 351–369.

FRACHOT, A. AND J-P LESNE (1993), Econometrics of Linear Factor Models of Interest Rates, working paper, Banque de France.

FRACHOT, A., J.P. LESNE, AND E. RENAULT (1995), Indirect Inference Estimation of Factor Models of the Yield Curve, Working Paper, Banque de France.

FRÜHWIRTH-SCHNATTER, S. AND A.L.J. GEYER (1997), Bayesian Estimation of Econometric Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure of Interest Rates via MCMC Methods, working paper, Vienna University of Economics and Business Administration.

GEYER, A.L.J. AND S. PICHLER (1997), A State-Space Approach to Estimate and Test Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure, working paper, University of Economics, Vienna.

GIBBONS, M.R. AND K. RAMASWAMY (1993), A Test of The Cox, Ingersoll, and Ross Model of the Term Structure, *Review of Financial Studies*, **6**, 619–658.

GONG, F. AND E.M. REMOLONA (1996), A Three Factor Econometric Model of the US Term Structure, Working Paper, Federal Reserve Bank New York.

GOURIEROUX, C. AND O. SCAILLET (1994), Estimation of the Term Structure from Bond Data, CREST working paper 9415.

HEATH, D., R. JARROW, AND A. MORTON (1992), Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation, *Econometrica*, **60**, 77–105.

HSIAO, C. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge.

HULL, J. AND A. WHITE (1990), Pricing Interest Rate Derivative Securities, *Review of Financial Studies*, **3**, 573–592.

HULL, J.C. (1996), *Options, Futures, and other Derivatives*, Prentice Hall, 3^{rd} edition.

JACQUIER, E. AND R.A. JARROW (1998), Dynamic Evaluation of Contingent Claim Model Error, *Journal of Econometrics*, Forthcoming.

JARROW, R.A. (1996), *Modelling Fixed Income Securities and Interest Rate Options*, McGraw-Hill.

KAPPI, J. (1997), Testing Affine Yield Factor Models, *Journal of Financial Engineering*, **6**, ??

LUND, J. (1997), Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates, Working Paper, Aarhus Business School.

MCCULLOCH, J.H. (1975), The Tax-Adjusted Yield Curve, *Journal of Finance*, **30**, 811–830.

NELSON, C.R. AND A.F. SIEGEL (1987), Parsimonious Modeling of Yield Curves, *Journal of Business*, **60**, 473–489.

RENAULT, E. (1996), Econometric Models of Option Pricing Errors, in D.M. Kreps and K.F. Wallis, (eds.), *Advance in Economics and Econometrics: Theory and Applications*, Cambridge University Press.

SCHOTMAN, P.C. (1996), A Bayesian Approach to the Empirical Valuation of Bond Options, *Journal of Econometrics*, **75**, 183–216.

SERCU, P. AND X. WU (1997), The Information Content in Bond Model Residuals: An Empirical Study on the Belgian Bond Market, *Journal of Banking and Finance*, **21**, 685–720.

VASICEK, O. (1977), An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, **5**, 177–188.

Table 1: Monte Carlo Results Vasicek Yield Curves

	RIMSE	r_1	r_2	r_3	κ	σ^2	θ
True values:							
	0.060	4	8	12	0.04	6.25	0
Fitted cross section models:							
Mean	0.001	4.00			0.066	4.33	0.92
Std.dev.	0.002	0.08			0.065	4.98	10.69
Minimum	0.000	3.68			-0.056	-42.14	-244.6
Maximum	0.020	4.28			0.336	7.97	7.96
Mean	0.001		8.00		0.100	1.80	1.35
Std.dev.	0.003		0.08		0.128	10.92	13.35
Minimum	0.000		7.77		-0.055	-149.6	-259.4
Maximum	0.033		8.27		0.682	10.44	10.55
Mean	0.004			12.00	0.201	-13.18	0.08
Std.dev.	0.008			0.08	0.351	45.41	36.42
Minimum	0.000			11.75	-0.360	-252.0	-1103.
Maximum	0.037			12.35	1.465	68.08	12.88
Pooled estimates:							
Mean	0.011	4.00	8.00	12.00	0.040	6.25	-0.29
Std.dev.	0.007	0.08	0.06	0.08	0.003	0.07	2.41
Minimum	0.000	3.67	7.82	11.73	0.028	6.02	-14.57
Maximum	0.038	4.25	8.20	12.36	0.054	6.56	6.17
True values:							
	0.200	4	8	12	0.04	6.25	0
Pooled estimates:							
Mean	0.053	4.00	8.00	12.01	0.040	6.19	-23.45
Std.dev.	0.039	0.40	0.31	0.39	0.016	0.62	100.11
Minimum	0.001	2.80	6.96	10.98	-0.013	-0.06	-1280.
Maximum	2.355	5.40	9.15	13.50	0.087	7.79	483.7

Notes: Yield curves are generated using the Vasicek model with the parameter values listed under "True values" including measurement noise:

$$e_t(\tau) = u_{1t} + \frac{\tau}{10}(u_{2t} - u_{1t})$$

where u_{1t} and u_{2t} are mutually uncorrelated normal random variables with zero mean. For the first part of the table the error standard deviations are 7 and 2 basispoints for u_{1t} and u_{2t} respectively. For the high noise case these standard deviations are 40 and 20 basispoints. All level parameters (r_1t, r_{2t}, r_{3t}, θ) are in units of percent per annum. The variance σ^2 is in percent per annum squared. The Root Integrated Mean Squared Error (RIMSE) is in basispoints. Parameters have been estimated by nonlinear least squares using 120 monthly spaced points on the yield curve. The longest maturity is ten years, the shortest one month.

Table 2: Parameter Estimates

	Equilibrium			Extended		
	One Factor	Two factors		One Factor	Two factors	
κ	0.042 (0.005)	0.003 (0.001)	0.311 (0.010)	0.095 (0.002)	0.012 (0.004)	0.081 (0.013)
σ	0.321 (0.023)	0.290 (1.122)	0.797 (0.248)	0.507 (0.026)	0.210 (0.014)	0.189 (0.015)
θ	-8.74 (5.18)	-4539 (1003)				
ϕ	0.725 (0.026)	0.031 (0.045)		0.720 (0.001)	0.503 (0.005)	
d	0.763 (0.008)	0.746 (0.046)		0.748 (0.003)	0.794 (0.008)	
c	0.785 (0.024)	0.693 (0.004)		-0.314 (0.005)	-0.356 (0.027)	
ω	0.408	0.258		0.398	0.265	
$\ln L$	31364	32214		31305	32143	

Notes: The table reports the estimation results for four versions of the (extended) Vasicek model in a panel with fixed effects. The first half of the columns refer to the equilibrium model, the last half of the columns to the extended Vasicek model. The parameters κ, σ and θ denote the mean reversion, volatility and infinite yield. The parameters ω, ϕ and d define the covariance matrix of the cross sectional error term; c is the autocorrelation parameter in the error term. Time in measured in years, so that for example $\ln 2/\kappa$ measures the halflife in years. Other parameters, like θ and σ are converted to percent per year. Standard errors are in parentheses. $\ln L$ denotes the log likelihood.

Table 3: Misspecification Tests

	Equilibrium		Extended	
	One Factor	Two factors	One Factor	Two factors
LM	38.51*	2.76	9.96*	9.01
df	4	6	4	6

Notes: The table reports the Lagrange Multiplier test of the hypothesis $H_0 : \gamma = 0$ and $D = 0$ in the extended model

$$\begin{aligned} \tilde{A}_t(\tau) &= A_t(\tau) + \gamma' g(\tau) \\ \tilde{B}(\tau) &= B(\tau) + Dg(\tau) , \end{aligned}$$

where $g(\tau) = (\tau, \tau^2)$ is a vector with two elements. The number of restrictions is given by df. An asterisk denotes significance at the 5% level.

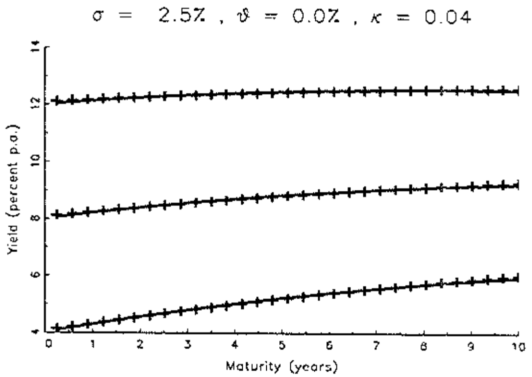

Figure 1: Yield Curve Pooling
The figure shows three yield curves generated by the Vasicek model with parameters $\theta = 0$, $\sigma = 2.5\%$ and $\kappa = 0.04$. The curves start at spot rates of $r = 4\%$, 8% and 12%, respectively.

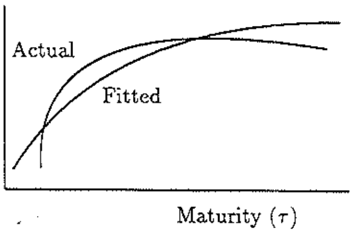

Figure 2: Yield curve fitting
The figure shows a hypothetical observed yield curve (Actual) and a yield curve implied by an equilibrium term structure model (Fitted). The observed yiled curve is drawn as a continuous line, because it is assumed that is constructed by interpolation methods like a spline function.

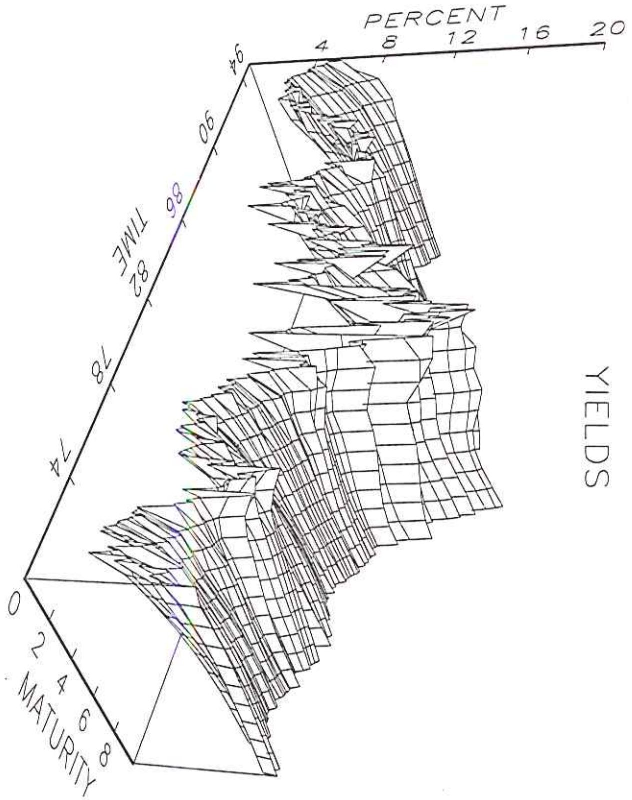

Figure 3: US Interest Rates

The figure shows monthly US discount yields for the period 1970 - 1994, and maturities between 0 and 10 years.

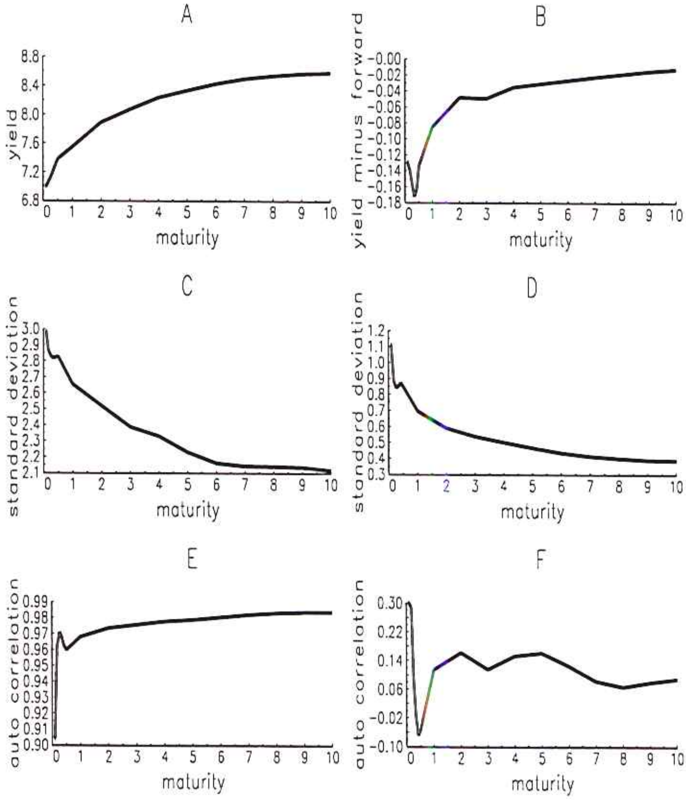

Figure 4: Data Summary

Panel A shows the sample mean of yields $Y_t(\tau)$ for different maturities, panel C the standard deviation, and panel E the first order autocorrelation. Yields are measured in percent per annum. Panels B, D en F provide the same summary for yields in deviation of the lagged forward rate $F_{t-1,t}(\tau)$ (solid line) and first differences of yields (dashed line). All data are monthly observations for the period 1970-1994.

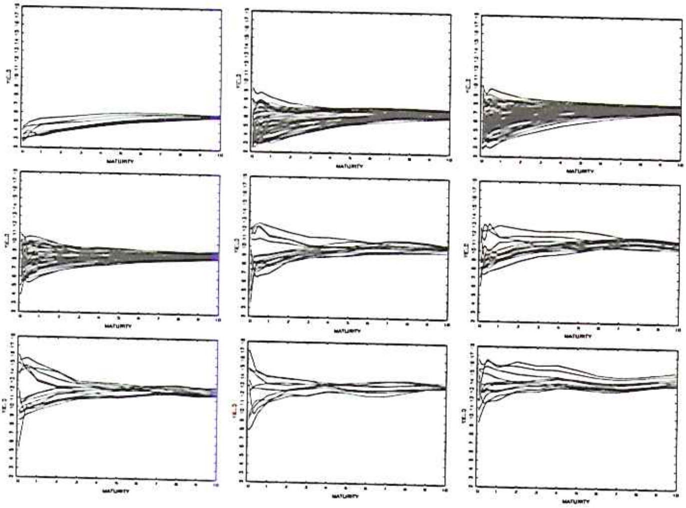

Figure 5: A second factor

The figure shows observed yield curves ordered with respect to the level of the ten year discount rate.

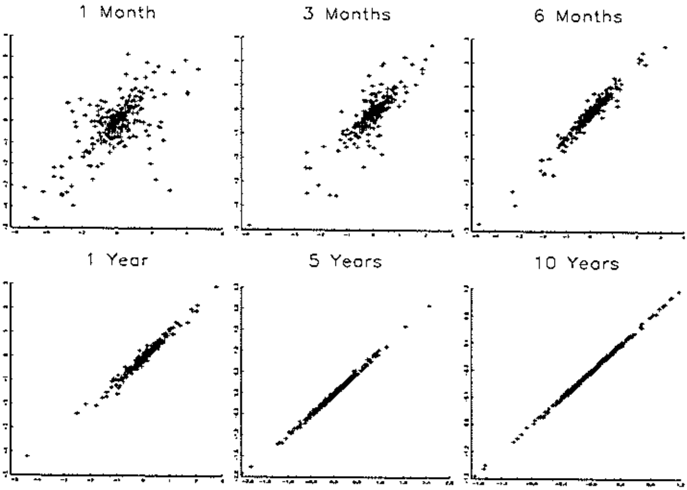

Figure 6: Discount yields and forward rates

The figure shows scatter plots with $Y_t(\tau) - F_{t,t-1}(\tau)$ on the vertical axis and $Y_t(\tau) - Y_{t-1}(\tau)$ on the horizontal axis for selected values of τ.

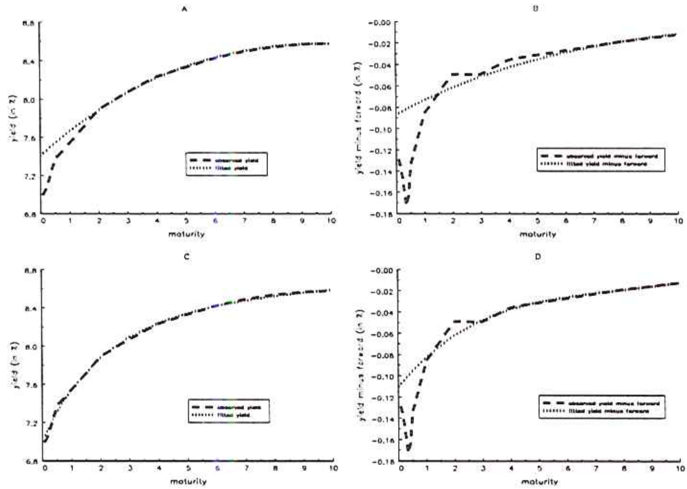

Figure 7: Average Yield Curve

This figure shows the average yield curve along with the fitted average yield curve for four different models: One factor equilibrium Vasicek (A), One factor extended Vasicek (B), Two factor equilibrium Vasicek (C), and Two factor extended Vasicek (D).

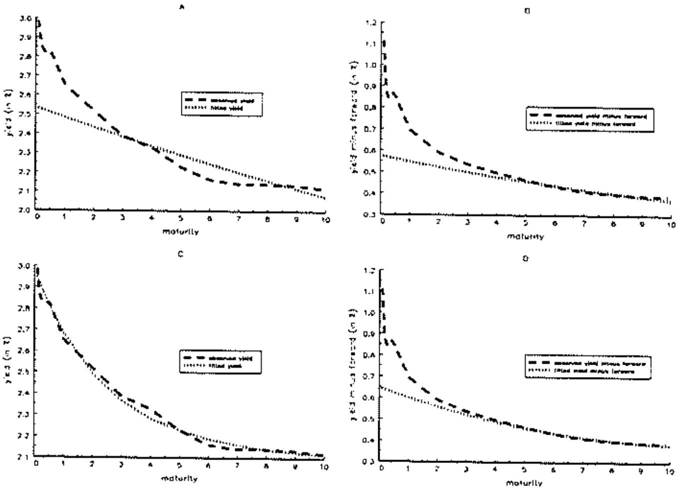

Figure 8: Term Structure of Volatilities

This figure shows the standard deviations of yields with different maturities along with the fitted standard deviations for four different models: One factor equilibrium Vasicek (A), One factor extended Vasicek (B), Two factor equilibrium Vasicek (C), and Two factor extended Vasicek (D). The fitted strandard deviation refer to the volatility of $B(\tau)x_t$. Units are percent per annum.

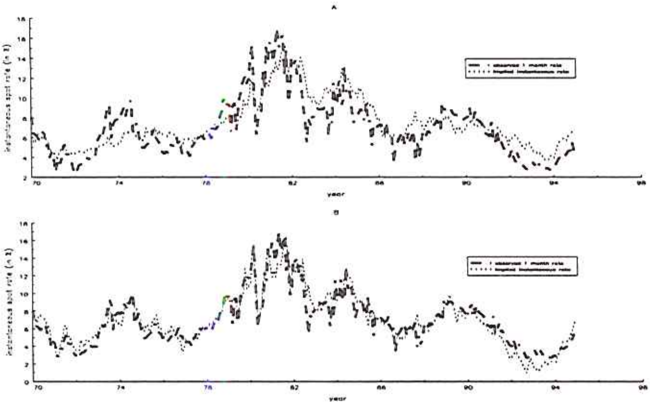

Figure 9: Implied spot rates

The figure shows the observed one month discount yield and the estimated spot rates r_t for the equilinbrium Vasicek model. Panel A refers to a one factor model, and panel B to the two factor model, where R_t is the sum of the two factors.

DISCUSSION PAPER SUBSCRIPTION FORM

Subscriptions may be placed for all CEPR Discussion Papers or for those appearing under one or more of the Centre's seven research programme areas: International Macroeconomics, International Trade, Industrial Organization, Financial Economics, Labour Economics, Public Policy and Transition Economics.

The quarterly charge will be determined by the number of papers sent during the preceeding three months and payment will be due on 31 March, 30 June, 30 September and 31 December of each year. New subscriptions must start from one of these dates. If no starting date is specified, the subscription will be started from the beginning of the next period. Papers are charged at the rate of £4 ($6). Individual academics may obtain papers at the concessionary rate of £3 ($4.50). To qualify for this concession, the declaration below (*) must be signed.

I wish to place a subscription for:

- ☐ Financial Economics (FE) Discussion Papers (c. 30 papers per year)
- ☐ Industrial Organization (IO) Discussion Papers (c. 25 papers per year)
- ☐ International Macroeconomics (IM) Discussion Papers (c. 100 papers per year)
- ☐ International Trade (IT) Discussion Papers (c. 40 papers per year)
- ☐ Labour Economics (LE) Discussion Papers (c. 25 papers per year)
- ☐ Public Policy (PP) Discussion Papers (c. 25 papers per year)
- ☐ Transition Economics (TE) Discussion Papers (c. 20 papers per year)

☐ * I wish to take advantage of the concessionary rate for individual academics. I am affiliated to an academic institution and will pay by personal cheque or credit card.

I want my subscription to start:

☐ 1 January ☐ 1 April
☐ 1 July ☐ 1 October

Back copies of papers from number 850 are available. For more details and information on out of print papers contact the Centre.

Name: ....

Position: ....

Email Address: ....

Telephone: ....

Fax: ....

Delivery Address: ....

....

....

☐ *Please invoice me each quarter*

Please invoice me in: ☐ sterling ☐ dollars ☐ euros

☐ *Please charge my Visa/Mastercard each quarter*

Credit Card No: ....

Expiry Date: ....

Cardholders' Name: ....

Signature: .... *Date:* ....

Please complete this form and return to:
The Subscription Officer
CEPR, 90–98 Goswell Road, London EC1V 7RR, UK
Fax: (44 171) 878 2999; Email: orders@cepr.org