Working Paper no. 291

Affine term structure models for the foreign exchange risk premium

Luca Benati

March 2006

Bank of England

Affine term structure models for the foreign exchange risk premium

*Luca Benati**

Working Paper no. 291

* Bank of England, Threadneedle Street, London, EC2R 8AH.
Email: luca.benati@bankofengland.co.uk.

The views expressed in this paper are those of the author, and not necessarily those of the Bank of England. I wish to thank Nikos Panigirtzoglou for helpful discussions, Liuren Wu for most helpful suggestions at an early stage of the project, Nicky Anderson, Peter Andrews, Mark Astley, Qiang Dai, participants at a seminar at the Bank of England, and at the C.V. Starr/*Review of Economic Dynamics* Conference 'Finance and the Macroeconomy', New York University, 11-12 October 2002, for comments, and David Backus for kindly providing the data set used in Backus, Telmer and Wu (1999). Usual disclaimers apply. This paper was finalised on 29 November 2005.

The Bank of England's working paper series is externally refereed.

Information on the Bank's working paper series can be found at
www.bankofengland.co.uk/publications/workingpapers/index.htm.

Publications Group, Bank of England, Threadneedle Street, London, EC2R 8AH;
telephone +44 (0)20 7601 4030, fax +44 (0)20 7601 3298, email mapublications@bankofengland.co.uk.

ISSN 1749-9135 (on-line)

Contents

Abstract 3

Summary 4

1 Introduction 6

2 A theoretical framework 9

3 The data 14

4 Some stylised facts 15

5 Empirical results 16

6 Conclusions 23

Tables and charts 24

References 35

Abstract

This paper uses two affine term structure models from the Duffie-Kan class—a three-factor Cox-Ingersoll-Ross model, and a three-factor model in the spirit of Longstaff and Schwartz—to extract historical estimates of foreign exchange risk premia for the pound with respect to the US dollar. The term structures of interest rates for the two countries are estimated jointly, together with the dynamics of the nominal exchange rates between them, *via* maximum likelihood. The likelihood function is computed *via* the Kalman filter, and is maximised numerically with respect to unknown parameters. Particular attention is paid to the robustness of the results across models; to the overall (filter plus parameter) econometric uncertainty associated with risk premia estimates; and to the ability of estimated structures to replicate Fama's 'forward discount anomaly'. The paper's main results may be summarised as follows. First, risk premia estimates are not consistent across the two models. Second, both models fail to replicate the forward discount anomaly, with theoretical values of β in the Fama regressions implied by estimated structures being consistently positive at all horizons from 1 to 12 months.

Key words: Foreign exchange risk premium; Fama puzzle; Duffie-Kan class; Kalman filter.

JEL classification: E30; E32

Summary

The ability to produce reliable estimates of foreign exchange risk premia would be of potentially paramount importance for policymakers. For example, a given appreciation of the currency bears markedly different implications for monetary policy when it originates from a movement in the risk premium, as opposed to (say) a change in the equilibrium exchange rate. Four decades ago, Fama first called the attention of the economic profession to the so-called 'forward discount anomaly', a puzzling violation of the uncovered interest parity (UIP) hypothesis according to which future foreign exchange rate depreciation should *exactly* reflect the current spread between foreign and domestic interest rates. Given that the presence of a time-varying foreign exchange risk premium represents a possible explanation for the failure of UIP to hold, in the intervening years economists have been trying to estimate risk premia within several different econometric frameworks. A first strand of literature has tried to estimate models based on strong theoretical restrictions, encountering, as of today, near-universal lack of success. Typical problems found within this approach include implausible estimates of the degree of risk aversion and, almost always, the empirical rejection of key theoretical implications of the underlying model.

A second group of studies has reacted to the rejection of models based on strong theoretical restrictions by pursuing a radically alternative strategy, namely by adopting a pure time-series approach that imposes a minimal theoretical structure on the data. While studies in this vein are capable of identifying a predictable component in the foreign exchange excess return, they typically suffer from the drawback that, by not imposing enough structure on the data, they cannot guarantee that such an estimated predictable component truly is a risk premium.

In this paper we adopt an intermediate approach, based on semi-structural models imposing minimal restrictions on the two countries' so-called pricing kernels — the processes on which all of the assets within the two countries, and the nominal exchange rate between them, can be priced. Such models should be considered as a 'bridge' between the two previously discussed groups of studies, imposing on a time-series structure a set of restrictions just sufficient to identify a foreign exchange risk premium with a reasonable degree of confidence, but otherwise leaving the model largely unconstrained. Although, on strictly logical grounds, it is clearly sub-optimal — ideally, we would like to be able to impose a solid theoretical structure capable of generating a time-varying risk premium — at the moment such an approach is probably the most promising.

We extract historical estimates of foreign exchange risk premia for the pound with respect to the US dollar based on two affine (ie, linear) term structure models. The term structures of interest rates for the two countries are estimated jointly, together with the dynamics of the nominal exchange rates between them, *via* maximum likelihood. The likelihood function is computed *via* the Kalman filter, and is maximised with respect to unknown parameters. Particular attention is paid to the robustness of the results across models; to the overall (filter plus parameter) econometric uncertainty associated with risk premia estimates; and to the ability of estimated structures to replicate Fama's 'forward discount anomaly', the key conditional stylised fact pertaining to the foreign exchange market.

The paper's main results may be summarised as follows. First, the risk premia estimates generated by the two models, although exhibiting a qualitatively similar time profile, are numerically quite different, to the point of casting doubts about the possibility of using them within a policy context. Second, both models fail to replicate the forward discount anomaly. Third — and not surprisingly, given the well-known difficulty of forecasting exchange rates — the estimated models exhibit virtually no forecasting power for foreign exchange rate depreciation.

1 Introduction

The ability to produce reliable estimates of foreign exchange risk premia would be of potentially paramount importance for policymakers. For example, an appreciation of the currency bears markedly different implications for monetary policy when it originates from a movement in the risk premium, as opposed to (say) a change in the equilibrium exchange rate. Since Fama (1984) first called the attention of the economic profession to puzzling violations of the uncovered interest parity (UIP) hypothesis (1)—for which the presence of a time-varying risk premium represents a possible explanation—economists have been trying to estimate foreign exchange risk premia within a variety of alternative econometric frameworks. A first strand of literature has tried to implement econometrically models based on strong theoretical restrictions (2)—coming, for example, from Lucas (1982)-type models. As of today the vast majority of these studies have been unsuccessful. Typical problems encountered in this literature are 'incredible' estimates of the risk aversion coefficient (3)—it is not uncommon to find estimates in excess of 50, or even of 100—and, in the majority of cases, the rejection of the overidentifying restrictions suggested by the underlying theory. (4)

A second group of studies has reacted to the rejection of models based on strong theoretical restrictions by pursuing a radically different strategy, namely by adopting a pure time-series approach that imposes a minimal structure on the data. (5) While studies in this vein are capable of identifying a predictable component in the foreign exchange excess return, they typically suffer from the drawback that, by not imposing enough structure on the data, they cannot guarantee that such an estimated predictable component truly represents a risk premium. Indeed, as stressed by Engel (1996),

---

(1) Specifically, *negative* values in the regression of subsequent nominal exchange rate depreciation on the forward discount—the so-called 'Fama puzzle'—instead of the unitary value predicted by the rational expectations hypothesis in the absence of a risk premium.

(2) See for example Mark (1985), Domowitz and Hakkio (1985), Hodrick (1989), Kaminsky and Peruga (1990), and Backus, Gregory and Telmer (1993).

(3) Mark (1985), for example, obtains estimates of the risk aversion coefficient ranging between 12.7 and 44.9. Hodrick (1989), using the dollar as the base currency, obtains an estimate of 60.9. Modjtahedi's (1991) estimates range between 6.5 and 64.9. Backus, Gregory and Telmer (1993) obtain, depending on the specification, either 52.8 or 107.1.

(4) Mark (1985), for example, rejects or does not reject the overidentifying restrictions depending on the specific set of instruments used. Both Modjtahedi (1991) and Backus, Gregory and Telmer (1993) reject the overidentifying restrictions, while Hodrick (1989) cannot reject them. An exception is the recent work of Groen and Balakrishnan (2005), for which econometric tests indicate that the model is not rejected by the data.

(5) See, for example, Cheung (1993), Taylor (1988), Hai, Mark and Wu (1997), Canova and Ito (1991).

> [. . .] a pure time series study of [the predictable component of the foreign exchange excess return] provides no evidence that [such a component] is a measure of a risk premium.

Given the current ‘state of the art’—ie, given the absence of a robust theoretical structure capable of generating a sizable foreign exchange risk premium, which is not consistently rejected by the data—the safest approach, at least for the time being, is probably to resort to semi-structural models imposing restrictions on the two countries’ pricing kernels. This type of model should, in a sense, be considered as a ‘bridge’ between the two previously discussed groups of studies, imposing on a time-series structure a set of restrictions which is just sufficient to identify with a reasonable degree of confidence a foreign exchange risk premium, but otherwise leaving the model largely unconstrained. Although, on strictly logical grounds, clearly suboptimal—ideally, we would like to be able to impose on the data a solid theoretical structure capable of generating a time-varying risk premium—at the moment such an approach is probably the most promising.

This paper uses affine multifactor models from the Duffie-Kan (1996, henceforth DK) class to extract historical estimates of foreign exchange risk premia for the pound with respect to the US dollar. The foreign exchange risk premium is modelled as an affine function of a vector of unobserved state variables, which are then extracted via Kalman filtering techniques.(6) There are two reasons for focusing on the DK class. First, it is currently the best-understood, having been completely described by the work of Duffie and Kan, and, as a result of this it is, as of today, the dominant one. Although other approaches have been, and are currently being developed,(7) the vast majority of recent studies of bond pricing have focused on the DK class.(8) Second, as shown by Backus, Foresi and Telmer (1996, 2001), the DK class is capable—at least, in principle—of replicating Fama’s ‘forward discount anomaly’,(9) thus allowing for the extraction of estimates of foreign exchange risk premia generated by a theoretical structure capable of reproducing all of the main moments of the data.

The paper is organised as follows. Section 2 discusses the theoretical framework underlying the

(6) The closely related work of Brandt and Santa-Clara (2002) jointly estimates interest rates dynamics within two countries, and the dynamics of the nominal exchange rate between them, via a simulated maximum likelihood estimator.

(7) See, for example, the quadratic class of models proposed by Leippold and Wu (2002, 2003).

(8) See, for example, Dai and Singleton (2000), Backus and Zin (1994), Backus, Foresi and Telmer (1996, 2001), and Backus, Telmer and Wu (1999).

(9) It is to be noticed, however, that the DK class of models is not the only one capable of replicating the Fama puzzle. Leippold and Wu’s quadratic class, for example, can also replicate the anomaly.

present study, starting with a brief exposition of no-arbitrage asset pricing theory, and then describing the main features of the DK class of exponentially affine multifactor models. Particular attention is paid to the ability of models belonging to the DK class to replicate the forward discount anomaly, a feature which, as shown by Backus, Foresi and Telmer (1996, 2001) crucially hinges on the presence of (at least) a common state variable exerting an asymmetric impact on the two countries' pricing kernels. Section 3 illustrates the data set used in the present study. The model is estimated by using both bond yields, and spot exchange rates,[(10)] as within the theoretical framework adopted herein, bond yields and spot exchange rates are driven by the very same stochastic processes—the two countries' pricing kernels. Section 4 reports some stylised facts for both bond yields and currency prices. In particular, the data clearly suggest the presence, in the term structures of interest rates of the two countries, of a common 'long' factor—ie of a factor exerting its influence mainly at the long end of the two yield curves. In Section 5 I report results from estimating two models, a three-factor Cox-Ingersoll-Ross (1985, henceforth, CIR) model, and a three-factor model in the spirit of Longstaff and Schwartz (1992), in which the log pricing kernel for each country is modelled as an affine function of three state variables: a common, long CIR factor; a country-specific short factor; and its conditional volatility. Section 6 concludes, and outlines possible directions for future research. A direction which appears to be particularly worth pursuing is, in the spirit of the recent work of Ang and Piazzesi (2001), to combine observed macroeconomic variables and latent factors within a no-arbitrage framework. As Ang and Piazzesi (2001) show, macroeconomic variables—in particular, inflation, and a measure of real activity—appear to be particularly important in explaining the dynamics of the short end of the yield curve—the one largely dominated by monetary policy actions—while latent factors dominate the long end of the curve, and still account for the vast majority of the overall variance.

(10) The key reason for restricting our attention to bond prices and foreign exchange rates is that these are the only assets whose prices are *uniquely* determined by the pricing kernel. To put it differently, as elaborated in Section 2.1 below, (a) knowledge of a country's pricing kernel is sufficient to uniquely determine that country's bond prices and bond yields; and (b) knowledge of two countries' pricing kernels is sufficient to uniquely determine the rate of change of the nominal exchange rate between them. On the other hand, for any other asset in the economy—for example, stock prices—knowledge of the pricing kernel is a necessary but not sufficient condition to determine its price—in the case of stock prices, for example, it is necessary to further specify a stochastic process for dividends.

2 A theoretical framework

2.1 Asset pricing theory

A well-known result from modern asset pricing theory is that in any arbitrage-free environment there exists[(11)] a positive random variable m_t—called the pricing kernel—satisfying

$$1 = E_t\left(m_{t+1}R_{t+1}\right) \qquad \textbf{(1)}$$

where R_{t+1} is the one-period nominal rate of return on an asset traded at time t. The importance of relationship **(1)** lies in its simplicity and in its generality: under the minimal assumption of no-arbitrage, theory guarantees the existence of the pricing kernel, which can then be used to price any kind of asset in the economy. In particular, it can easily be shown that, first, assuming knowledge of the stochastic properties of a country's pricing kernel, it is straightforward to derive the prices of bonds at all maturities, and therefore both nominal interest rates and forward rates at all horizons.[(12)] Second, assuming knowledge of the stochastic properties of two countries' pricing kernels, the dynamics of the nominal exchange rate between them can then be trivially determined: Backus, Foresi and Telmer (2001) indeed show that, given equation **(1)** for the home country, and the analogous relationship

$$1 = E_t\left(\tilde{m}_{t+1}\tilde{R}_{t+1}\right) \qquad \textbf{(2)}$$

for the foreign country, the following relationship[(13)] holds

$$s_{t+1} - s_t = \ln \tilde{m}_{t+1} - \ln m_{t+1} \qquad \textbf{(3)}$$

(where s_t is the logarithm of the nominal exchange rate, defined as the price of a unit of foreign currency expressed in units of domestic currency), ie nominal exchange rate depreciation is equal to the difference between the logarithms of the two pricing kernels. Finally, since it can easily be shown that the forward premium is equal to

$$f_t - s_t = \ln \tilde{m}_{t+1|t} - \ln m_{t+1|t} \qquad \textbf{(4)}$$

(where f_t is the logarithm of the forward nominal exchange rate, and $|t$ indicates the expectation conditional on information available at time t), it immediately follows that the foreign exchange risk premium—defined as the 'wedge' between the forward rate and the expected spot rate—is

(11) In particular, the *existence* of a (not necessarily unique) pricing kernel is guaranteed by the absence of arbitrage opportunities, while its *uniqueness* requires the additional assumption of market completeness. (On this, see for example the discussion in Backus, Foresi and Telmer (2001).

(12) See, for example, Backus, Foresi and Telmer (1998), and Backus, Telmer and Wu (1999).

(13) Equation **(3)** follows from a simple arbitrage condition on the foreign exchange market.

given by

$$\rho_t \equiv f_t - s_{t+1|t} = \ln \tilde{m}_{t+1|t} - \ln m_{t+1|t} + E_t\left(\ln m_{t+1}\right) - E_t\left(\ln \tilde{m}_{t+1}\right) \qquad \textbf{(5)}$$

As stressed by Backus *et al* (2001), the symmetry of expression **(5)** suggests a possible reason for the overall failure of ARCH and GARCH-in-mean models of the risk premium—see for example, Domowitz and Hakkio (1985), and Bekaert and Hodrick (1993):

> [o]ne view of this failure is that GARCH-M models violate our sense of symmetry: an increase in the conditional variance of the depreciation rate increases risk on both sides of the market, and hence carries no presumption in favor of one currency or the other. [...] GARCH-M models, to put it simply, focus on the wrong conditional variance. (14)

Assuming, further, that the two countries' log pricing kernels are conditionally normally distributed—as is routinely done in the literature—namely

$$\ln m_{t+1}|I_t \sim N\left(\mu_{m,t}, \sigma^2_{m,t}\right) \qquad \textbf{(6)}$$

$$\ln \tilde{m}_{t+1}|I_t \sim N\left(\mu_{\tilde{m},t}, \sigma^2_{\tilde{m},t}\right) \qquad \textbf{(7)}$$

(where I_t is the information set available at time t), it can easily be shown that expected nominal exchange rate depreciation uniquely depends on the conditional means of the two log kernels, namely

$$s_{t+1|t} - s_t = \mu_{\tilde{m},t} - \mu_{m,t} \qquad \textbf{(8)}$$

while the foreign exchange risk premium uniquely depends on their conditional volatilities: (15)

$$\rho_t = \frac{\sigma^2_{\tilde{m},t} - \sigma^2_{m,t}}{2} \qquad \textbf{(9)}$$

(In related work, Brandt and Santa-Clara (2002) and Brandt, Cochrane and Santa-Clara (2005) show that the volatility of exchange rates has important information about the discount factors of the two countries, making it interesting to use second moments of the exchange rate in any empirical exercise.)

The preceding discussion suggests that, if we were able to extract from the data reasonably precise estimates of the stochastic processes followed by the two countries' pricing kernels, getting historical estimates of the foreign exchange risk premium would become, from a strictly technical

(14) To put it crudely, it is not clear why an increase in the volatility of the dollar/sterling exchange rate should make the dollar more attractive.
(15) The derivation of expressions **(8)** and **(9)** exploits the well-known property that, if a variable X is normally distributed with mean μ_X and variance σ^2_X, exp(X) is lognormal, and E[exp(X)]= exp[μ_X+0.5σ^2_X].

point of view, a trivial task. Furthermore, the fact that the two countries' pricing kernels drive *both* exchange rate dynamics, *and* the dynamics of bond prices (interest rates) in the two countries, suggests that—at first sight quite paradoxically—the best strategy to estimate the foreign exchange risk premium is to exploit the information contained in the two countries' term structures—as is well known, exchange rate changes are very close to white noise, so that they contain virtually no information.

2.2 The Duffie-Kan (1996) class of affine multifactor models

Duffie and Kan (1996) provide a complete characterisation of the so-called exponential affine—affine, for short—class of models, in which log bond prices and bond yields at the various maturities are affine functions of a vector of (possibly unobserved) state variables, showing how several well-known bond pricing models—among them, the classic Vasicek (1977), Brennan and Schwartz (1979), Longstaff and Schwartz (1992), and Cox, Ingersoll and Ross (1985) models—represent particular cases of such a class. Following Backus, Foresi and Telmer's (2001) rendition in discrete time of DK's original continuous-time analysis, the DK class is described by the following two equations:

$$-\ln m_{t+1} = \delta + \gamma' z_t + \lambda' V(z_t)^{\frac{1}{2}} \epsilon_{t+1} \quad \textbf{(10)}$$

$$z_{t+1} = (I_k - \Psi)\,\Xi + \Psi z_t + V(z_t)^{\frac{1}{2}} \epsilon_{t+1} \quad \textbf{(11)}$$

where m_t is the pricing kernel, δ is a scalar, γ is a $k \times 1$ vector of constants (which can be interpreted as the 'loadings' of the state variables onto the pricing kernel), $'$ stands for transposition, λ is a $k \times 1$ vector of constants, z_t is a $k \times 1$ vector of state variables evolving according to **(11)**, Ψ is a stable matrix with positive diagonal elements, I_k is the $k \times k$ identity matrix, Ξ is the vector of the unconditional means for the state variables, and $V(z_t)$ is a diagonal matrix capturing time variation in the volatility structure, with typical element

$$\upsilon_i(z_t) = \alpha_i + \beta_i' z_t \quad \textbf{(12)}$$

Finally, a set of additional restrictions on the parameters space is necessary in order to ensure that the state vector z_t never leaves the region defined by non-negative values of the volatility functions υ_i (see Appendix A of Backus, Foresi and Telmer (2001), or DK (1996).

Given such a structure, by defining as b_t the market price, as of time *t*, of a bond of maturity *n*—ie a claim to one pound at time *t*+*n* in all possible states of the world—and by applying the relationship

$$b_t^{n+1} = E_t \left(m_{t+1} b_{t+1}^n \right) \tag{13}$$

which holds by definition, it can be easily shown that minus log bond prices are given by

$$-\ln b_t^n = A_n + B_n' z_t \tag{14}$$

which immediately implies the following expression for bond yields as functions of the state vector z_t [16]

$$y_t = n^{-1} \left[A_n + B_n' z_t \right] \tag{15}$$

where A_n, a scalar, and B_n, a $k \times 1$ vector, evolving according to

$$\begin{bmatrix} A_{n+1} \\ B_{n+1} \end{bmatrix} = \begin{bmatrix} \delta \\ \gamma \end{bmatrix} + \begin{bmatrix} 1 & \Xi' \left(I_k - \Psi \right)' \\ 0_k & \Psi' \end{bmatrix} \begin{bmatrix} A_n \\ B_n \end{bmatrix} -$$

$$-\frac{1}{2} \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_k \\ \beta_1 & \beta_2 & \dots & \beta_k \end{bmatrix} \begin{bmatrix} (\lambda_1 + B_{1n})^2 \\ (\lambda_2 + B_{2n})^2 \\ \dots \\ (\lambda_k + B_{kn})^2 \end{bmatrix} \tag{16}$$

with initial conditions[17] A_0=0, and B_0=0_k, with 0_k being a $k \times 1$ vector of zeros.

2.3 Accounting for the Fama puzzle within the Duffie-Kan class

Backus, Foresi and Telmer (2001) discuss how affine models belonging to the DK class are—at least in principle—capable of replicating the Fama puzzle by allowing for an asymmetric impact of the vector of common state variables on the two countries' pricing kernels.[18] Specifically, consider the structure **(10)-(11)**, and assume that the foreign country's pricing kernel is given by the foreign equivalent of **(11)**, ie by

$$-\ln \tilde{m}_{t+1} = \tilde{\delta} + \tilde{\gamma}' z_t + \tilde{\lambda}' V \left(z_t \right)^{\frac{1}{2}} \epsilon_{t+1} \tag{17}$$

(16) Bond yields are defined as y_t^n=-$n^{-1}\ln b_t^n$.
(17) The initial conditions are an immediate consequence of the fact that $b_t^0 \equiv 1$—ie the price of one pound today is one pound.
(18) As shown by Backus, Foresi and Telmer (2001), in principle these models can replicate the Fama puzzle under an alternative circumstance, namely when the two countries' pricing kernels depend both on two vectors of idiosyncratic (ie country-specific) factors, and on a vector of common factors exerting an identical influence on the two kernels. They demonstrate, however, how a necessary condition for such a structure to be able to replicate the forward discount anomaly is to allow nominal interest rates to take negative values with a strictly positive probability. To put it differently, a DK model in which (a) interest rates cannot take negative values with a positive probability, and the two countries' pricing kernels depend either (b) uniquely on two vectors of idiosyncratic factors, or (c) both on two vectors of idiosyncratic factors, and on a vector of common factors exerting a symmetric impact on the two kernels, is in principle incapable of replicating the anomaly. Given the logical problems associated with allowing nominal interest rates to take negative values, we have ignored such a possibility.

Equations **(10)**, **(11)** and **(17)** imply that[19] the one-period-ahead depreciation rate and the one-period-ahead forward discount are respectively given by

$$s_{t+1} - s_t = \left(\delta - \tilde{\delta}\right) + (\gamma - \tilde{\gamma})' z_t + \left(\lambda - \tilde{\lambda}\right)' V(z_t)^{\frac{1}{2}} \epsilon_{t+1} \tag{18}$$

$$f_t - s_t = r_t - \tilde{r}_t =$$

$$= \left[\left(\delta - \tilde{\delta}\right) - \frac{1}{2}\sum_{j=1}^{k} \alpha_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right)\right] + \left[\left(\lambda - \tilde{\lambda}\right) - \frac{1}{2}\sum_{j=1}^{k} \beta_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right)\right]' z_t \tag{19}$$

which, in turn, implies that the theoretical value of β in the Fama regression

$$s_{t+1} - s_t = \alpha + \beta (f_t - s_t) + v_{t+1} \tag{20}$$

is equal to

$$\beta = \frac{\text{Cov}(s_{t+1} - s_t, f_t - s_t)}{\text{Var}(f_t - s_t)} = \frac{\left[(\gamma - \tilde{\gamma}) - \frac{1}{2}\sum_{j=1}^{k} \beta_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right)\right]' \text{Var}(z_t) (\gamma - \tilde{\gamma})}{\text{Var}(f_t - s_t)} \tag{21}$$

Expression **(21)** clearly illustrates how the ability of the structure **(10)-(11)-(17)** to replicate the forward discount anomaly—ie a negative estimate of β in **(20)**—crucially depends on the two quantities $(\gamma - \tilde{\gamma})$ and $\sum_{j=1}^{k} \beta_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right)$—in other words, it depends on the existence of asymmetric effects of z_t on the two countries' pricing kernels, either 'directly', through the γ's (the 'loading factors' of the state variables onto the pricing kernels), or 'indirectly', through the vectors of prices of risk (the λ's). On the other hand, a comparison of **(21)** with the expression for the risk premium,

$$\rho_t = -\frac{1}{2}\sum_{j=1}^{k} \alpha_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right) - \frac{1}{2}\left[\sum_{j=1}^{k} \beta_j \left(\lambda_j^2 - \tilde{\lambda}_j^2\right)\right]' z_t \tag{22}$$

clearly shows that the sign of the estimates from the Fama regression bears no immediate connection to the sign and extent of the foreign exchange risk premium, so that results from the Fama regressions cannot be used to draw conclusions on the existence and extent of foreign exchange risk premia, in the specific sense that there is no one-to-one mapping between results from the Fama regressions and the extent of the foreign exchange risk premium.

Given the scant attractiveness of a model in which nominal interest rates are allowed to take negative values, in this paper we have decided to pursue the second avenue discussed by Backus, Foresi and Telmer (2001), adopting a model in which a common state variable—which, in what follows, we interpret as a 'long' factor, ie as a factor affecting bond yields mainly at the long end of the curve—is allowed to exert an asymmetric impact on the two countries' log pricing kernels.

(19) See Backus, Foresi and Telmer (1996, Section 5.3).

3 The data

The models used in this paper have been estimated based on a data set comprising US bond yields from the data set used by Backus, Telmer and Wu (1999), and UK bond yields from the Bank of England database. Bond yields from the Bank of England database have been constructed via the 'variable roughness penalty' (VRP) spline curve method described in Anderson and Sleath (2001). Yields from the Backus *et al* (1999) data set, on the other hand, have been constructed via the smoothed Fama-Bliss method. For reasons of methodological homogeneity, we would have preferred to use, both for the United Kingdom and for the United States, bond yields from the Bank of England database only. Unfortunately, such a database extends back to the beginning of the 1980s only for the United Kingdom, while for the United States yield curves based on the VRP method are available only starting from 1992. Estimating the two models described below based on UK and US bond yields from the Bank of England database over the period July 1992-May 2002 turned out to be basically infeasible: the convergence properties of the maximum likelihood algorithm were very poor, and final estimates were quite imprecise. We therefore decided to employ a longer data set comprising the Bank of England and Backus *et al* (1999) data. From a strictly technical point of view, the Fama-Bliss and VRP methods are quite similar. They are both based on splines, but while the Fama-Bliss method first estimates the splines and then smoothes, the VRP method estimates the splines and smoothes at the same time.

Table A provides a comparison between US bond yields from the two data sets for the period of overlapping, from July 1992 to December 2000. As the table makes clear, the difference is not especially marked. In particular, for maturities between six months and ten years—the ones used in estimation in this paper—the difference varies between an average of 1.28 basis points (with a standard deviation of 6.19 basis points) at the six-month maturity, and an average of 3.45 basis points (with a standard deviation of 6.96 basis points) at the ten-year maturity. Overall, the gain from having a much longer data set to work with largely offsets the drawback originating from the lack of methodological homogeneity in the construction of the data.[20] The spot foreign exchange

(20) A second minor problem in mixing the Bank of England and Backus *et al* (1999) data sets is that for eight observations, out of an overall length of the sample of 240 monthly observations, the day on which the UK yields from the Bank of England data set, and the US yields from the Backus *et al* data set, were sampled is not the same. The Backus *et al* data set has consistently been sampled on the last working day of each month. In constructing the Bank of England data set for UK yields (which is based on original daily observations) we tried to match perfectly the dates of the Backus data set, but for eight observations this was not possible, and we chose to take the closest available observation. In two cases the difference is four days, while in the remaining cases it is three days. (Another possibility would have been to treat these observations as missing, but we preferred to keep them.)

data for the pound *vis-à-vis* the US dollar are from Datastream. The sample period is from January 1980 to December 2004. The Backus *et al* (1999) data set for the United States does not have any missing observation. As for the United Kingdom, the Bank of England database has a few missing observations at the very short end of the curve.[21] Given that, within the theoretical framework adopted herein, all the assets are driven by the same vector of state variables, a partial solution to such a problem is to expand the cross-sectional dimension of the data set, with the inclusion of a relatively large number of maturities.

4 Some stylised facts

4.1 Evidence on the existence of a common 'long' factor in international term structures

Tables B and C, and Chart 1, provide evidence on the existence of a common 'long' factor in UK and US term structures—ie a factor mainly affecting the long ends of the two bond yield curves. Table B shows the fractions of variance explained by the first four static principal components[22] extracted from first-differenced 18, 24, 36, 48, 60, 72, 96, and 120-month UK and US bond yields.[23] As the table clearly shows, the first static principal component explains exactly two thirds of the overall variance within the matrix of first-differenced bond yields, thus clearly suggesting the existence of a common factor in the two countries' term structures. The next question is then: which portions of the two bond yield curves are most closely correlated? Chart 1, plotting demeaned and standardised 9-month, and 1, 2, 4, 6, and 10-year UK and US bond yields, suggests that—in line with conventional wisdom—the correlation is stronger at the longer maturities. This is indeed the case: the contemporaneous correlation between first-differenced UK and US bond yields rises monotonically from 0.34 for the 18-month maturity, to almost 0.48 for

(21) The use of the Kalman filter to compute the likelihood provides an ideal way of dealing with the presence of missing observations. On this, see for example Watson and Engle (1983).

(22) Given a $T \times K$ matrix of K covariance-stationary series of length T, the first N static principal components are orthogonal linear combinations of the K columns explaining, in decreasing order, the greatest amount of variance within the matrix.

(23) The reason for considering the first difference of bond yields, instead of their levels, is to take into account of the possible presence of unit roots, which cannot be rejected at conventional levels based on standard tests. It is important to stress however, how, from a strictly conceptual point of view, the notion of a unit root in interest rates is manifestly nonsensical. Nominal rates are indeed equal to the sum of the Wicksellian natural rate of interest, and of expected inflation. The Wicksellian rate quite obviously cannot contain a unit root. And if the central bank acts in a purposeful way, and targets a constant rate of inflation, inflation must necessarily be mean-reverting, thus implying that, under rational expectations, expected inflation must be mean-reverting too. But this implies that nominal rates also cannot contain a unit root. (The apparent non-stationarity of nominal rates over the sample period might, quite obviously, be the consequence of a small sample largely dominated by the episode of high inflation of the 1970s. A longer sample would most likely capture mean-reversion in the level of nominal rates.)

the 10-year yields, thus suggesting that the common factor mainly exerts its influence at the very long and of the two countries' bond yield curves.[24] This suggests the adoption of models with both a common 'long' factor, and country-specific 'short' factors, ie factors exerting their influence mainly at the short end of the yield curves.

4.2 Results from Fama regressions

Table D reports results for the Fama regression **(20)** for the UK pound *vis-a-vis* the US dollar at four different horizons, one, three, six, and twelve months. The sample period is January 1980-December 2004.[25] Estimates of β in equation **(20)** are in line with existing, well-known empirical evidence, with all of the estimates being consistently negative at all horizons. As previously mentioned in Section 2, the importance of the results from the Fama regressions lies in the fact that any well-specified model of foreign exchange rate determination, and any credible candidate model for estimating the foreign exchange risk premium, must be capable of replicating these crucial conditional moments of the data.

5 Empirical results

5.1 A three-factor CIR model

In the spirit of Hodrick and Vassalou (2002), we start by considering a three-factor CIR model.[26] Specifically, for each country the log pricing kernel is assumed to be an affine function of a common CIR factor affecting mainly the long end of the yield curve (which is allowed to exert an asymmetric impact on the countries' kernels), and of two country-specific CIR factors. The model is therefore described by the following equations:

$$-\ln m_{j,t+1} = \left(\gamma_j + \frac{\lambda_{C,j}^2}{2}\right) z_{C,t} + \left(1 + \frac{\lambda_{1,j}^2}{2}\right) z_{1,j,t} + \left(1 + \frac{\lambda_{2,j}^2}{2}\right) z_{2,j,t} + $$
$$+ \lambda_{C,j} z_{C,t}^{\frac{1}{2}} \epsilon_{C,t+1} + \lambda_{1,j} z_{1,j,t}^{\frac{1}{2}} \epsilon_{1,j,t+1} + \lambda_{2,j} z_{2,j,t}^{\frac{1}{2}} \epsilon_{2,j,t+1} \quad \text{for } j = UK, US \qquad \mathbf{(23)}$$

(24) On the other hand, due to the presence of missing observations at the very short end of the UK curve, results for both the nine-month and the one-year maturity have been computed for the two subsamples indicated in the table. Results are markedly different across subsamples, with the period between April 1982 and November 1989 characterised by a significantly higher correlation than the latter period.

(25) Results are based on the Richard Levich data set (at http://bertha.gsia.cmu.edu/telmerc/misc.html), available for the period January 1973-December 1994, which we updated based on Datastream.

(26) Results from a two-factor CIR model (contained in a previous version of the paper) are qualitatively similar to the ones presented herein, but slightly inferior in terms of the fit of the term structures of interest rates for the two countries. They are however available upon request.

$$z_{C,t+1} = \mu_C \left(1 - \phi_C\right) + \phi_C z_{C,t} + \sigma_C z_{C,t}^{\frac{1}{2}} \epsilon_{C,t+1} \quad \textbf{(24)}$$

$$z_{h,k,t+1} = \mu_{h,k} \left(1 - \phi_{h,k}\right) + \phi_{h,k} z_{h,k,t} + \sigma_{h,k} z_{h,k,t}^{\frac{1}{2}} \epsilon_{h,k,t+1} \quad \text{for } k = UK, US, k = UK, US \quad \textbf{(25)}$$

where the notation is obvious, and C indicates the common factor. It can be easily shown that the expression for the foreign exchange risk premium is given by

$$\rho_t^{UK,US} = \frac{z_{C,t}\left(\lambda_{C,US}^2 - \lambda_{C,UK}^2\right) + \lambda_{1,US}^2 z_{1,US,t} + \lambda_{2,US}^2 z_{2,US,t}}{2} - \frac{\lambda_{1,UK}^2 z_{1,UK,t} + \lambda_{2,UK}^2 z_{2,UK,t}}{2} \quad \textbf{(26)}$$

—the foreign exchange risk premium is therefore a linear function of the UK and foreign country-specific factors, and of the common 'long' factor—while the theoretical value of β in the Fama regression **(20)** is equal to:

$$\beta_{UK,US} = 1 + \frac{1}{2} \frac{\left(\gamma_{UK} - \gamma_{US}\right)\left(\lambda_{UK,C}^2 - \lambda_{US,C}^2\right) \frac{\mu_C \sigma_C^2}{1-\phi_C^2} + \sum_{j=1}^2 \left[\lambda_{j,UK}^2 \frac{\mu_{j,UK}\sigma_{j,UK}^2}{1-\phi_{j,UK}^2} + \lambda_{j,US}^2 \frac{\mu_{j,US}\sigma_{j,US}^2}{1-\phi_{j,US}^2}\right]}{\left(\gamma_{UK} - \gamma_{US}\right)^2 \frac{\mu_C \sigma_C^2}{1-\phi_C^2} + \sum_{j=1}^2 \left[\lambda_{j,UK}^2 \frac{\mu_{j,UK}\sigma_{j,UK}^2}{1-\phi_{j,UK}^2} + \lambda_{j,US}^2 \frac{\mu_{j,US}\sigma_{j,US}^2}{1-\phi_{j,US}^2}\right]} \quad \textbf{(27)}$$

From **(27)** it is immediately apparent that values for the γ_j's (the loading factors) different from one are necessary in order to allow the model to replicate the forward discount anomaly—more precisely, if all of the γ_j's are equal to one, the model is implicitly imposing a theoretical value of β in the Fama regression greater than one.[(27)] Since for any bilateral rate only one of the two γ_j's need to be different from one, in what follows we set γ_{UK} equal to one, and we estimate γ_{US}. Finally, due to a well-known identification problem typical of this class of models—the difficulty in separately econometrically identifying more than one 'level' parameter for each term structure[(28)]—we set $\mu_{2,UK}$=$\mu_{1,US}$=$\mu_{2,US}$=10^{-3} (setting them equal to zero caused convergence problems with the maximum likelihood algorithm). As model **(23)-(25)** belongs to the DK class, bond yields can be trivially computed via the formulas **(15)-(16)**. The model can then be cast in state-space form, with observation and transition equations given by **(15)** and **(11)** respectively, and can be estimated via maximum likelihood, by computing the log-likelihood via the Kalman filter, and maximising it numerically with respect to unknown parameters—for details, see eg Hamilton (1994, chapter 13). A technical problem in computing the log-likelihood is that since the factors, which are unobserved, act as their own volatilities, an exact likelihood function is, strictly speaking, impossible to compute, because the covariance matrices are unknown. In the spirit of Harvey Ruiz, and Sentana (1992), and following Kim and Nelson (2000, section 6.1.3), we therefore replaced the unknown covariance matrices in the Kalman filtering algorithm with their

(27) Given that the μ's are obviously all positive.
(28) See, eg, Backus, Telmer and Wu (1999, page 10).

estimates conditional on information at time t-1, and we computed an approximated log-likelihood via the resulting approximated Kalman filter.[29]

Table E reports maximum likelihood estimates of the model's structural parameters, together with estimated standard errors (in parentheses).[30] Optimisation was performed by means of the MATLAB subroutine `fminsearch.m`, based on the Nelder-Mead simplex algorithm.[31] The observation equation included the one-month nominal exchange rate depreciation (ie the first difference of the log exchange rate), and, for each of the two countries, the 6, 12, 18, 24, 36, 60, 72, and 120-month maturities of nominal interest rates. As for the transition equation, it was simply given by **(24)-(25)**, cast in the matrix form **(11)**. Chart 2 shows two-sided estimates of the factors, together with the UK and US 18-month and 10-year bond yields (both factors and bond yields have been demeaned and standardised); the UK and US actual average term structures of interest rates, together with the upper and lower 90% theoretical confidence bands generated by the estimated model;[32] and the two-sided estimated foreign exchange risk premium, together with the 90% confidence bands, computed via the Hamilton (1985) Monte Carlo procedure to take into account of both filter and parameter uncertainty.[33]

As expected, the common factor is significantly correlated with the long ends of the two countries' bond yield curves, and is very persistent. As for the country-specific factors, both the first UK factor and the first US factor are very strongly correlated with the UK 18-month bond yield, and respectively with the US 18-month yield, and are both very persistent, with estimated autoregressive parameters close to 1. The second UK factor and the second US factor, on the other hand, are less persistent, with estimated autoregressive parameters around 0.3. As for the term structures of interest rates, the UK actual term structure falls entirely within the 90% theoretical confidence intervals generated by the estimated model, while in the case of the United States the actual average term structure falls slightly outside the theoretical confidence bands both at the very

(29) The formulas for the approximated Kalman filtering algorithm (contained in a previous version of the paper, but not reported here) are available upon request.
(30) Standard errors have been obtained by inverting the estimated information matrix, computed via the Berndt, Hall, Hall and Hausman (1974) 'outer product' formula.
(31) In estimation I imposed the restriction that the autoregressive parameters be smaller than one.
(32) Theoretical confidence bands have been computed via Monte Carlo, based on 10,000 replications. For each replication, I drew, for each single parameter, from a normal distribution with mean equal to the parameter's MLE estimate, and with a standard deviation equal to the parameter's estimated standard error (reported in Table E).
(33) Parameter uncertainty originates from to the fact that the true values of the model's structural parameters are unknown, and ought to be estimated. Filter uncertainty, on the other hand, originates from the fact that factors are unobserved, and ought to be extracted via the Kalman filter (by definition, filter uncertainty would be there even if the model's structural parameters were known with certainty).

short and at the very long end of the curve. While, from a strictly conceptual point of view, this represents a rejection of the model, at the 90% confidence level, it is also important to stress that the rejection, as the chart makes clear, is very mild. The two-sided estimate of the UK-US foreign exchange risk premium mostly reflects movements in the common long factor, as $\lambda_{1,UK}$, $\lambda_{2,UK}$, $\lambda_{1,US}$, and $\lambda_{2,US}$ are all estimated to be close to zero. Both the filter and the overall econometric uncertainty associated with two-sided estimated foreign exchange risk premia are quite small, most likely reflecting the vast amount of information used in estimation. Our estimates imply a positive—although modest—risk premium on the pound over the entire sample period, with a peak of about 0.4-0.5 percentage points between the begining of the sample and mid-1980s, and a progressive decline over subsequent years. It is interesting to notice that the bulk of the decrease in the risk premium is estimated to have taken place around the time of the large fall in the dollar engineered by the Plaza Accord of 1985—in this sense, the comparatively large and positive risk premium on the pound of previous years appears, *ex post*, to have correctly signalled the possibility of sterling's appreciation *vis-à-vis* the dollar.

Chart 3 plots, for the two countries, actual and two-sided estimated bond yields at various maturities, while Chart 4 shows (top panels) estimated autocorrelations of UK and US two-sided pricing errors at various maturities, and the pound-dollar actual depreciation, together with the one-step-ahead forecast produced by the estimated model. The two-sided pricing errors are quite significantly autocorrelated, especially at the long end of the curve for the United Kingdom,[34] and at the short end of the curve for the United States. The difficulty of getting a low autocorrelation of the pricing errors is quite common in the literature. De Jong and Santa-Clara (1999), for example, based on a two-factor CIR model, report autocorrelations at lag one which, depending on the specific maturity, are between 0.45 and 0.77;[35] and the classic Pearson and Sun (1994) empirical implementation of a two-factor CIR model based on maximum likelihood techniques has an empirical performance which, as far as fitting actual term structures is concerned, is quite unsatisfactory (see in particular their Figure 2). So ours is, in a sense, a common problem in the literature, magnified by the fact that, in the present case, we are fitting two term structures at the same time.[36]

(34) As for the long end of the UK yield curve, institutional features of the gilt market may help explain the model's poor fit.
(35) Qualitatively similar results can be found in De Jong (2000).
(36) On the other hand, De Jong (2000), based on a three-factor model, obtains a remarkably good performance in terms of fitting the US term structure of interest rates.

Where the model dramatically fails is in replicating the results from the Fama regressions. Table G reports results from 10,000 stochastic simulations of the estimated model.(37) For each replication, we drew, for each single parameter, from a normal distribution with mean equal to the parameter's MLE estimate, and with a standard deviation equal to the parameter's estimated standard error. Based on the drawn parameters, we then generated artificial time series for the factors, and based on these we computed bond yields for the two countries, the forward discounts, and nominal exchange rate depreciation. Finally, based on the generated nominal exchange rate depreciation, and on the generated forward discounts at the various horizons, we ran Fama regressions for horizons from one to twelve months. As the table shows (a) the mean value of β based on the 10,000 replications increases monotonically from 1.13 at the one-month horizon to 12.14 at the twelve-month horizon, and (b) the 90% confidence interval never contains a negative value, and goes from [0.58; 1.70] at the one-month horizon to [4.11; 18.26] at the twelve-month horizon. As previously stressed, the importance of the results from the Fama regressions lies in the fact that any credible candidate estimate of the foreign exchange risk premium must be compatible with these crucial conditional moments of the data. Under this respect, the fact that the estimated model is incapable of generating a negative value of β in the Fama regression **(20)** clearly casts serious doubts on the reliability of the risk premia estimates reported in Chart 2.

5.2 A three-factor model in the spirit of Longstaff and Schwartz (1992)

We now consider a three-factor model in the spirit of Longstaff and Schwartz (1992).(38) Specifically, for each country the log pricing kernel is modelled as an affine function of three state variables: a common CIR factor affecting mainly the long end of the yield curve, which is allowed to exert an asymmetric impact on the countries' log kernels; a country-specific factor affecting mainly the short end of the curve; and its conditional volatility. For each country the dynamics of the logarithm of the pricing kernel is therefore governed by

$$-\ln m_{j,t+1} = \frac{\lambda_{j,V}^2 \sigma_{j,V}^2}{2} + \begin{bmatrix} 1 & \frac{\lambda_{j,\gamma}^2}{2} & \left(\kappa_j + \frac{\sigma_\theta^2 \lambda_{j,\theta}^2}{2}\right) \end{bmatrix} \begin{bmatrix} r_{j,t} \\ V_{j,t} \\ \theta_t \end{bmatrix} +$$

(37) In order to make the results exactly comparable to the ones reported in Table D, the model was re-estimated over the sample period January 1980-December 1994, and stochastic simulations were performed based on these estimates.
(38) In the original two-factor Longstaff-Schwartz (1992) model, the log pricing kernel is an affine function of the short rate and of its conditional volatility.

$$+\begin{bmatrix} \lambda_{j,\gamma} & \lambda_{j,V} & \lambda_{j,\theta} \end{bmatrix} \begin{bmatrix} V_{j,t}^{\frac{1}{2}} & 0 & 0 \\ 0 & \sigma_{j,V} & 0 \\ 0 & 0 & \sigma_{\theta}\theta_t^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \epsilon_{j,r,t+1} \\ \epsilon_{j,V,t+1} \\ \epsilon_{\theta,t+1} \end{bmatrix} \quad \textbf{(28)}$$

for $j = UK, US$, where the notation is obvious, and where the long, common factor, θ_t, the short, country-specific factors, $r_{j,t}$, and their conditional volatilities, $V_{j,t}$, evolve according to

$$\theta_{t+1}\text{-}\mu_{\theta} = \rho_{\theta}\left(\theta_t\text{-}\mu_{\theta}\right) + \sigma_{\theta,t}^{\frac{1}{2}}\epsilon_{\theta,t+1} \quad \textbf{(29)}$$

$$\begin{aligned} r_{j,t+1}\text{-}\mu_j &= \phi_j\left(r_{j,t}\text{-}\mu_j\right) + V_{j,t}^{\frac{1}{2}}\epsilon_{j,r,t+1} \\ V_{j,t+1}\text{-}\eta_j &= \rho_j\left(V_{j,t}\text{-}\eta_j\right) + \sigma_{j,V}\epsilon_{j,V,t+1} \end{aligned} \quad \textbf{(30)}$$

It can be easily shown that **(28)-(30)** imply the following theoretical value of $\beta_{UK,j}$ in the Fama regression **(20)**:

$$\beta_{UK,j} = 1 + \frac{1}{2}\frac{\left(\kappa_{UK}\text{-}\kappa_j\right)\sigma_{\theta}^2\left(\lambda_{UK,\theta}^2\text{-}\lambda_{j,\theta}^2\right)\frac{\mu_{\theta}\sigma_{\theta}^2}{1-\rho_{\theta}^2}}{\left(\kappa_{UK}\text{-}\kappa_j\right)^2\frac{\mu_{\theta}\sigma_{\theta}^2}{1-\rho_{\theta}^2} + \frac{\eta_{UK}}{1-\phi_{UK}^2} + \frac{\eta_j}{1-\phi_j^2}} \quad \textbf{(31)}$$

and the following expression for the foreign exchange risk premium:

$$\rho_t^{UK,j} = \text{-}\frac{\lambda_{UK,V}^2\sigma_{UK,V}^2\text{-}\lambda_{j,V}^2\sigma_{j,V}^2}{2}\text{-}\frac{\lambda_{UK,V}^2}{2}V_{UK,t}+\frac{\lambda_{j,V}^2}{2}V_{j,t} - \frac{\sigma_{\theta}^4}{2}\left(\lambda_{UK,\theta}^2 - \lambda_{j,\theta}^2\right)\theta_t \quad \textbf{(32)}$$

Expression **(32)** shows how, consistent with the discussion in Section 2, the foreign exchange risk premium uniquely depends on the conditional second moments of the log pricing kernel, here represented by the conditional volatilities of the two country-specific factors, and by the ‘long’ common CIR factor, which acts as its own volatility.

Table F reports maximum likelihood estimates of the model’s structural parameters. For the same reasons as in the previous sections, κ_{UK} and μ_{US} were set equal to 1 and 0 respectively. Again, both the common ‘long’ factor and the two country-specific factors were estimated to be highly persistent, with estimated autoregressive parameters close to 1, while the conditional volatilities of the two factors were estimated to be significantly less persistent, with autoregressive parameters around 0.6.

Chart 5 shows the estimated two-sided common factor, together with the UK and US 10-year bond yields (again, demeaned and standardised); the UK country-specific factor, together with the UK 18-month yield; the US country-specific factor, together with the US 18-month yield; the conditional volatilities of the two country-specific factors; the actual average term structures of interest rates, together with the upper and lower 90% theoretical confidence bands generated by the estimated model; and the two-sided estimated foreign exchange risk premium, together with the 90% confidence bands, again computed via the Hamilton (1985) Monte Carlo procedure. Again—and not surprisingly—the common long factor is significantly correlated with the long ends of the two countries' yield curves, and both the UK and US country-specific factors display a significant correlation with the short ends of the two countries' yield curves. Both the UK and the US actual average term structures of interest rates are fully inside the 90% theoretical confidence bands generated by the estimated model, but the confidence bands are so wide, reflecting substantial imprecision of the estimates, that it is not really clear what to make of this. The overall imprecision of the estimates is reflected in the comparatively large parameter uncertainty for the estimated risk premium, which is substantially greater than in the case of the three-factor CIR model.

As for the estimated risk premium, although the time profile is qualitatively similar to the profile of the estimate produced by the three-factor CIR model—with both estimates mostly reflecting movements in the estimated common long factor—the level is instead quite markedly higher. Again, the risk premium is estimated to have been comparatively large during the period preceding the Plaza Accord of 1985, and to have quite markedly declined since then. The fall in the risk premium around 1985, however, appears as less drastic than the one estimated based on the three-factor CIR model, so that the estimates based on the present model paint, overall, a picture of a gradual decline over the entire sample period.

Chart 6 plots, for the two countries, actual and two-sided estimated bond yields at various maturities, while Chart 7 (top panels) plots estimated autocorrelations of UK and US two-sided pricing errors at various maturities, and the pound-dollar actual depreciation, together with the one-step-ahead forecast produced by the estimated model. In terms of the autocorrelation of the two-sided pricing errors, the performance of the model is roughly comparable to that of the three-factor CIR model.

As in the case of the three-factor CIR model, the estimated structure fails to replicate the results from the Fama regressions. As in the previous section, the ability of the estimated model to replicate the results from the Fama regressions reported in Table G was assessed via stochastic simulations and, as in the previous section (a) the theoretical value of β in the Fama regression rises monotonically from 1.00 to 10.77, and (b) none of the 90% confidence intervals contains a negative value. As we stressed in the previous section, failure to replicate results from the Fama regressions casts doubts on the reliability of the risk premium estimates shown in Chart 5.

6 Conclusions

This paper has used two affine term structure models from the Duffie-Kan (1996) class—a three-factor Cox-Ingersoll-Ross (1985) model, and a three-factor model in the spirit of Longstaff and Schwartz (1992)—to extract historical estimates of foreign exchange risk premia for the pound with respect to the US dollar. The term structures of interest rates for the two countries have been estimated jointly, together with the dynamics of the nominal exchange rates between them, via maximum likelihood. The likelihood function has been computed via the Kalman filter, and has been maximised numerically with respect to unknown parameters. Particular attention has been paid to the robustness of the results across models; to the overall (filter plus parameter) econometric uncertainty associated with risk premia estimates; and to the ability of estimated structures to replicate Fama's (1984) 'forward discount anomaly'.

The paper's main results may be summarised as follows. First, the risk premia estimates generated by the two models, although exhibiting a qualitatively similar time profile, are numerically quite different, to the point of casting doubts about the possibility of using them within a policy context. Second, both models fail to replicate the forward discount anomaly, with theoretical values of β in the Fama regressions implied by estimated structures being consistently positive at all horizons from one to twelve months. Third—and not surprisingly, given the well-known difficulty of forecasting exchange rates—estimated models exhibit virtually no forecasting power for foreign exchange rate depreciation.

As for possible directions for future research, one that appears to be particularly worth pursuing is, in the spirit of the recent work of Ang and Piazzesi (2001), to combine observed macroeconomic variables and latent factors within a no-arbitrage framework. As Ang and Piazzesi (2001) show, macroeconomic variables—in particular, inflation and a measure of real economic activity—appear to be particularly important in explaining the dynamics of the short end of the yield curve, the one largely dominated by monetary policy actions, while latent factors dominate the long end of the curve, and still account for the vast majority of the overall variance.

Table A: US yields: a comparison between the Backus *et al* (1999) and the Bank of England data sets		
Mean and standard deviation of the difference between the two sets of yields (basis points)		
Maturity (in months)	Mean	Standard deviation
6	-1.28	6.19
12	-1.55	5.56
18	-1.91	4.77
24	-2.87	4.41
36	-1.75	4.32
60	-1.85	4.81
72	-3.14	5.28
120	-3.45	6.96

Table B: Fractions of variance explained by the first four static principal components extracted from first-differenced UK and US bond yields			
First	Second	Third	Fourth
0.666	0.282	0.031	0.014

Table C: Contemporaneous correlations between first-differenced UK and US bond yields at different maturities	
Maturity	*Correlation*
9-month (Apr. 1982-Nov. 1989)	0.368
9-month (Jan. 1991-Dec. 2000)	0.128
1-year (Jan. 1980-Nov. 1989)	0.350
1-year (Jan. 1990-Dec. 2000)	0.192
18-month	0.340
2-year	0.353
3-year	0.366
4-year	0.378
5-year	0.396
6-year	0.417
8-year	0.455
10-year	0.479
For details, see text.	

Table D: Results from Fama regressions				
	Horizon (in months)			
	1	3	6	12
α	-4.5E-3	-1.5E-2	-2.16E-2	-3.8E-2
	(2.2E-3)	(7.4E-3)	(0.012)	(0.02)
β	-1.09	-1.34	-0.67	-0.50
	(0.78)	(0.77)	(0.75)	(0.71)
Newey-West (1987) standard errors in parentheses. The sample period is January 1980-December 2004.				

Table E: Maximum likelihood estimates, three-factor CIR model

ϕ_θ	0.997 (3.58E-04)	$\mu_{1,UK}$	4.97E-03 (1.35E-04)	$\mu_{2,US}$	Set to 10^{-3}	$\lambda_{1,UK}$	-0.038 (8.48E-05)
$\phi_{1,UK}$	0.998 (2.09E-04)	σ_θ	4.30E-03 (1.71E-04)	γ_{UK}	Set to 1	$\lambda_{2,UK}$	0.068 (4.60E-04)
$\phi_{2,UK}$	0.291 (4.48E-03)	$\sigma_{1,UK}$	4.57E-03 (2.28E-04)	$\sigma_{1,US}$	6.41E-03 (5.06E-04)	$\lambda_{1,US}$	0.125 (4.00E-03)
$\phi_{1,US}$	0.988 (1.70E-03)	$\sigma_{2,UK}$	0.012 (9.76E-04)	$\sigma_{2,US}$	0.011 (2.59E-03)	$\lambda_{2,US}$	0.150 (3.52E-03)
$\phi_{2,US}$	0.320 (0.016)	$\mu_{2,UK}$	Set to 10^{-3}	$\lambda_{\theta,UK}$	-1.759 (0.018)	γ_{US}	2.503 (0.012)
μ_θ	1.78E-03 (1.21E-04)	$\mu_{1,US}$	Set to 10^{-3}	$\lambda_{\theta,US}$	-1.837 (3.56E-03)	σ_u	3.42E-04 (7.18E-04)

For details on estimation, see text.

Table F: Maximum likelihood estimates, three-factor model in the spirit of Longstaff and Schwartz (1992)

μ_θ	5.51E-03 (2.78E-05)	ρ_{UK}	0.613 (3.59E-05)	μ_{US}	Set to 0	$\lambda_{\theta,US}$	-45.265 (1.63E-05)
ρ_θ	0.996 (2.24E-06)	$\sigma_{r,UK}$	1.06E-05 (3.88E-06)	ϕ_{US}	0.978 (3.94E-05)	$\lambda_{V,US}$	Set to 0
σ_θ	1.96E-02 (5.13E-06)	$\lambda_{r,UK}$	-42.981 (1.00E-04)	η_{US}	1.03E-06 (1.29E-06)	κ_{US}	1.023 (2.50E-05)
μ_{UK}	1.52E-03 (2.69E-04)	$\lambda_{\theta,UK}$	-42.348 (1.23E-06)	ρ_{US}	0.602 (1.52E-04)	σ_u	6.52E-03 (4.31E-06)
ϕ_{UK}	0.955 (1.11E-07)	$\lambda_{V,UK}$	Set to 0	$\sigma_{r,UK}$	5.47E-06 (9.24E-08)		
η_{UK}	9.46E-07 (1.69E-06)	κ_{UK}	Set to 1	$\lambda_{r,US}$	-46.603 (2.46E-05)		

For details on estimation, see text.

Table G: Theoretical values of β in the Fama regressions implied by estimated models				
	Three-factor CIR model		*Three-factor model à-la Longstaff-Schwartz*	
Horizon (in months)	Mean value of β	90% confidence interval	Mean value of β	90% confidence interval
1	1.13	[0.58; 1.70]	1.00	[0.82; 1.23]
2	2.26	[1.09; 3.44]	1.99	[1.61; 2.43]
3	3.36	[1.57; 5.15]	2.96	[2.35; 3.63]
4	4.44	[1.97; 6.81]	3.91	[3.04; 4.80]
5	5.48	[2.35; 8.42]	4.83	[3.67; 5.98]
6	6.50	[2.72; 9.98]	5.74	[4.26; 7.17]
7	7.50	[3.00; 11.46]	6.64	[4.75; 8.36]
8	8.48	[3.28; 12.90]	7.51	[5.20; 9.61]
9	9.43	[3.54; 14.29]	8.35	[5.56; 10.85]
10	10.35	[3.74; 15.63]	9.18	[5.86; 12.09]
11	11.26	[3.93; 16.94]	9.99	[6.11; 13.36]
12	12.14	[4.11; 18.26]	10.77	[6.34; 14.63]
Results based on 10,000 stochastic simulations. For details, see text.				

Chart 1: Informal evidence on the existence of a common 'long' factor in UK and US bond yield curves: 9-month, 1, 2, 4, 6, and 10-year UK and US bond yields, demeaned and standardised

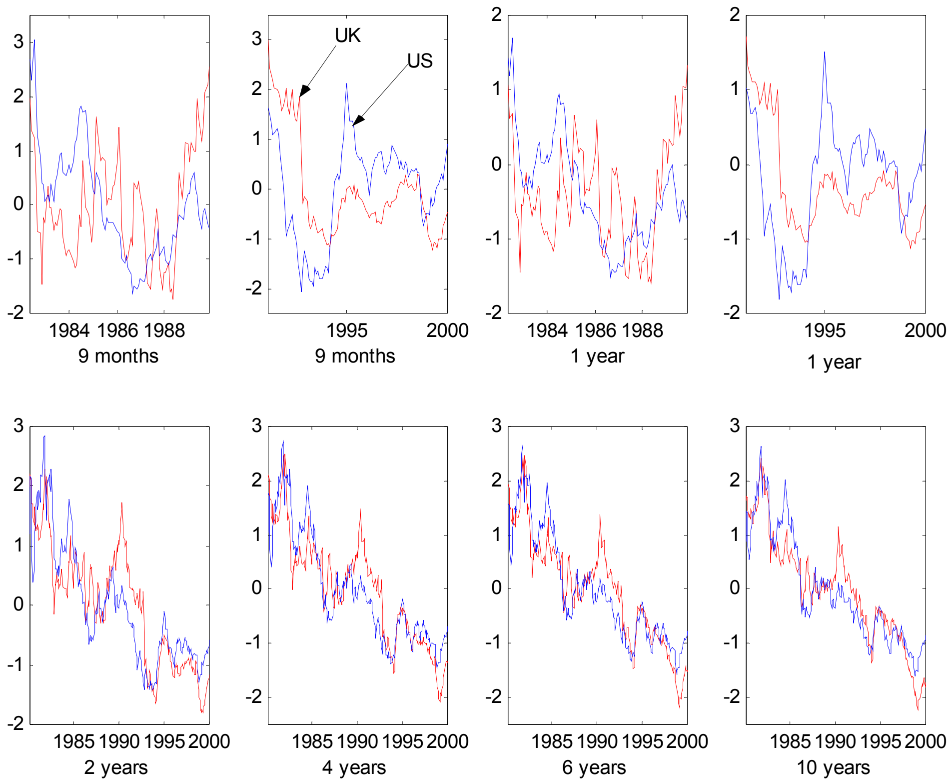


Chart 2: A three-factor CIR model for the UK and the US: two-sided estimates of the factors (factors and bond yields demeaned and standardised), actual and theoretical term structures, two-sided estimates of the foreign exchange risk premium, and 90% confidence bands (confidence bands, computed via Monte Carlo, take into account of both filter and parameter uncertainty; for technical details see text)

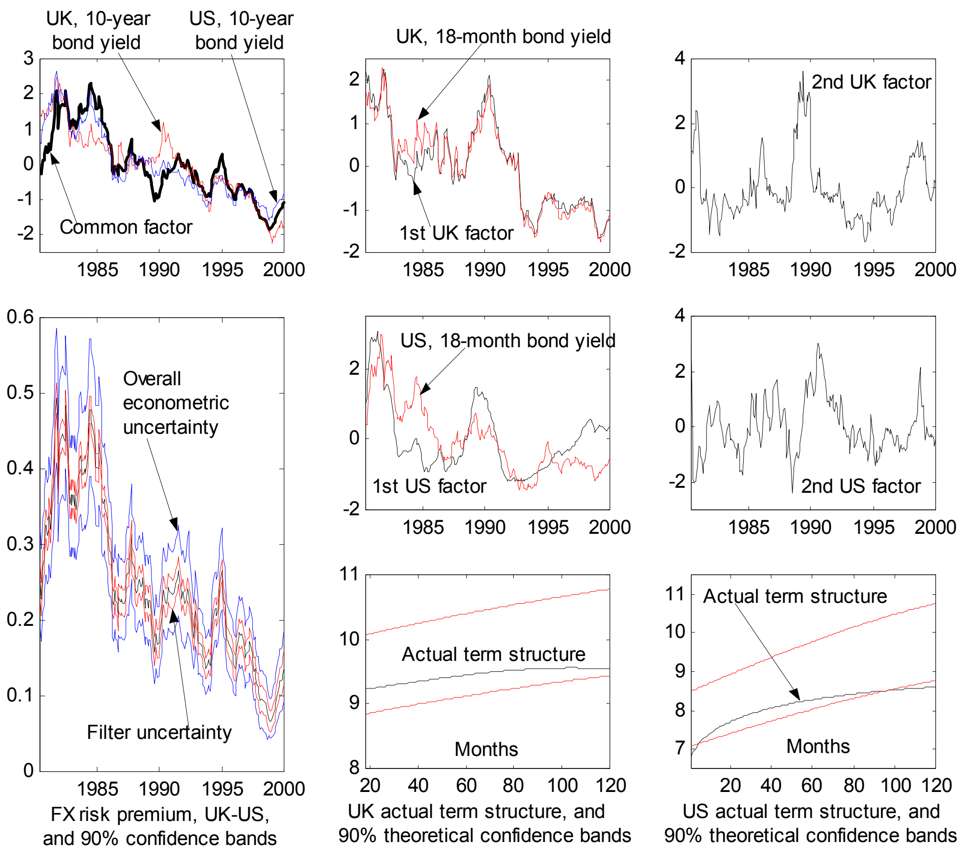


Chart 3: A three-factor CIR model for the UK and the US: actual and theoretical, two-sided bond yields

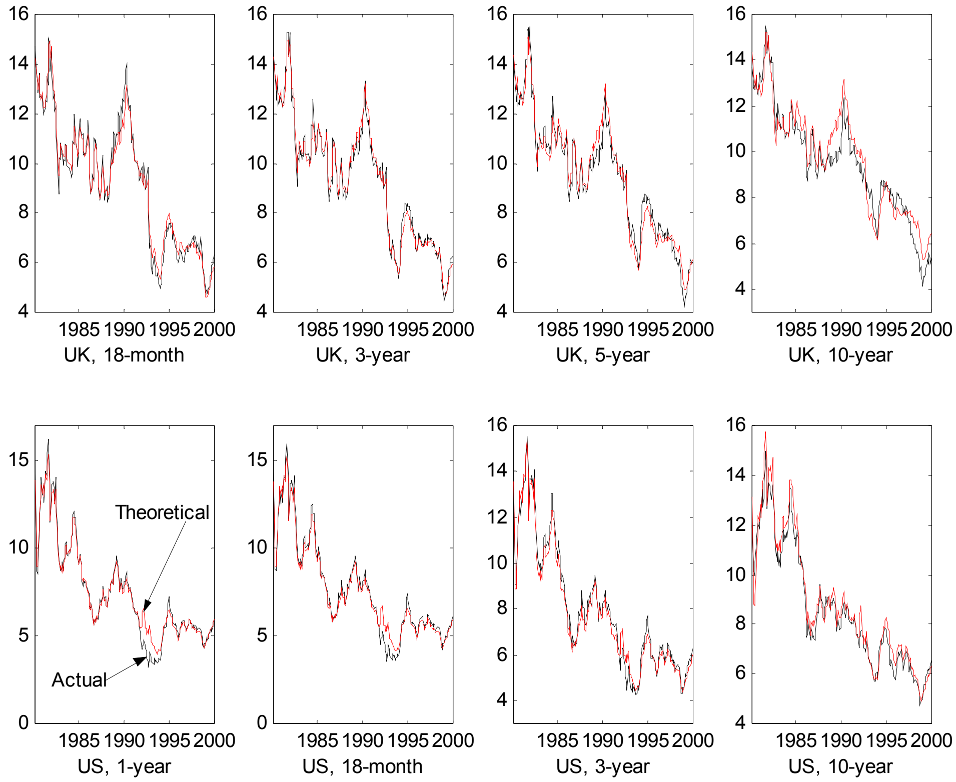


Chart 4: A three-factor CIR model for the UK and the US: autocorrelations of two-sided errors for bond yields, actual exchange rate depreciation, and one-step-ahead forecast

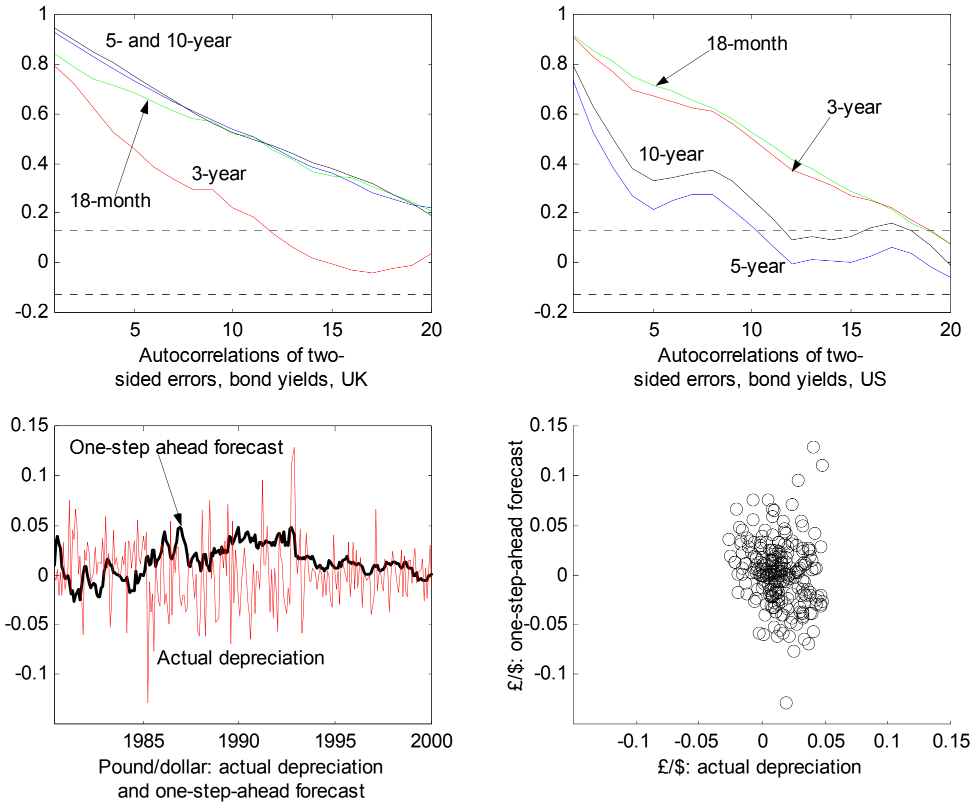


Chart 5: A three-factor model in the spirit of Longstaff and Schwartz (1992) for the UK and the US: estimated factors, estimated volatilities, and bond yields (factors and bond yields demeaned and standardised; plotted confidence bands only take into account the filter uncertainty; for technical details see text)

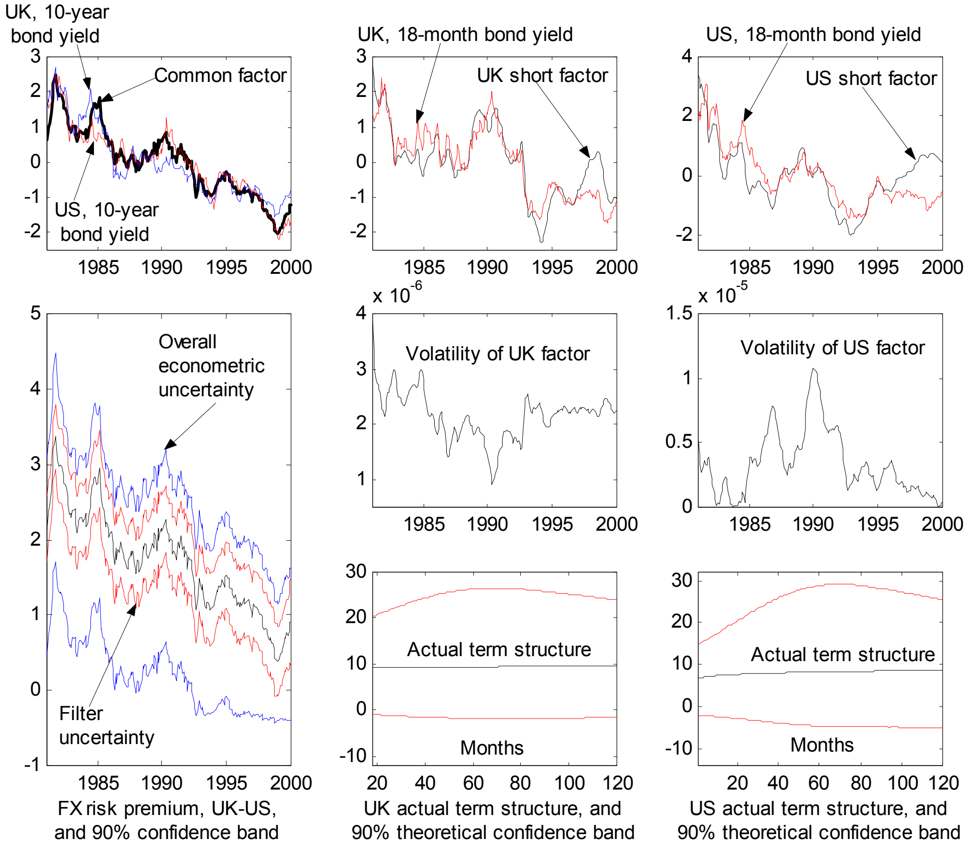


Chart 6: A three-factor model in the spirit of Longstaff and Schwartz (1992) for the UK and the US: actual and theoretical, two-sided bond yields

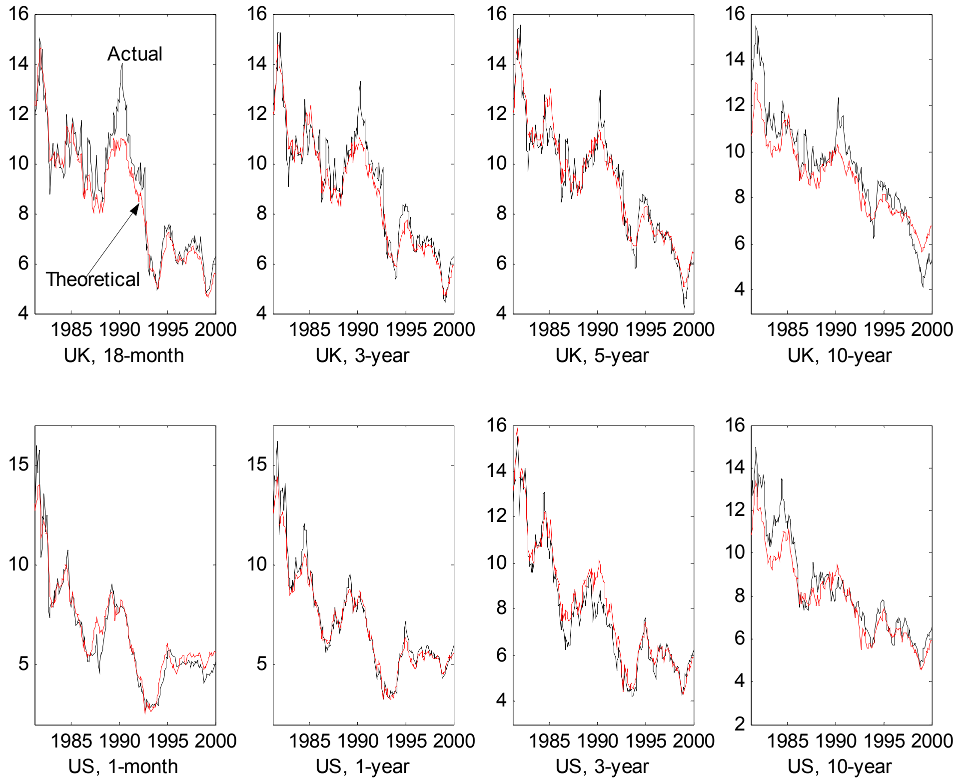
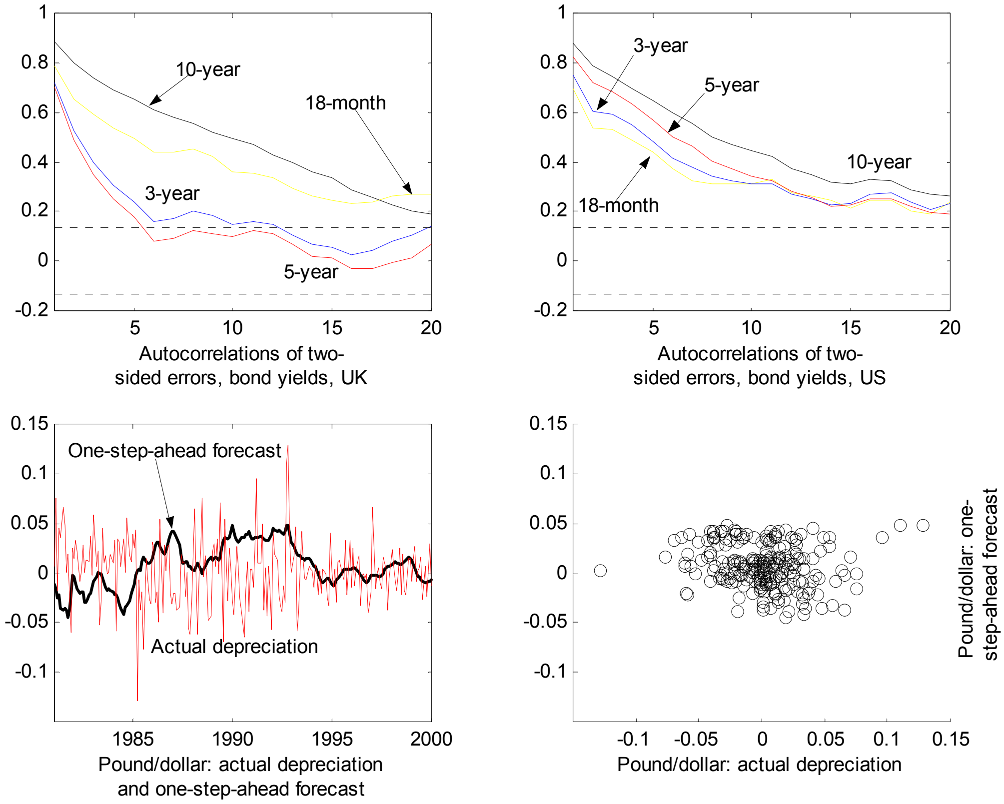


Chart 7: A three-factor model in the spirit of Longstaff and Schwartz (1992) for the UK and the US: actual and fitted (two-sided) exchange rate, one-step-ahead forecast error, actual one-month depreciation and one-step-ahead forecast, and autocorrelation functions of two-sided errors for bond yields and exchange rates



References

Anderson, N and Sleath, J (2001), 'New estimates of the UK real and nominal yield curves', *Bank of England Working Paper no. 126*.

Ang, A and Piazzesi, M (2001), 'A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables', *NBER working paper no. 8363.*

Backus, D, Foresi, S and Telmer, C (1996), 'Affine models of currency pricing', *NBER Working Paper no. 5623.*

Backus, D, Foresi, S and Telmer, C (1998), 'Discrete-time models of bond pricing', New York University, *mimeo*.

Backus, D, Foresi, S and Telmer, C (2001), 'Affine term structure models and the forward premium anomaly', *Journal of Finance*, Vol. 56, pages 279-304.

Backus, D, Gregory, A and Telmer, C (1993), 'Accounting for forward rates in markets for foreign currency', *Journal of Finance*, XLVIII, pages 1,887-908.

Backus, D, Telmer, C and Wu, L (1999), 'Design and estimation of affine yield models', New York University, *mimeo.*

Backus, D and Zin, S (1994), 'Reverse engineering the yield curve', *NBER working paper no. 4676.*

Bekaert, G and Hodrick, R J (1993), 'On biases in the measurement of foreign exchange risk premiums', *Journal of International Money and Finance*, Vol. 12, No. 2, pages 115-38.

Berndt, E, Hall, B, Hall, R and Hausman, J (1974), 'Estimation and inference in nonlinear structural models', *Annals of Economic and Social Measurement*, Vol. 3-4, pages 653-65.

Brandt, M and Santa-Clara, P (2002), 'Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets', *Journal of Financial Economics*, Vol. 63, pages 161–210.

Brandt, M, Cochrane, J and Santa-Clara, P (2005), 'International risk sharing is better than you think, or: exchange rates are too smooth', *Journal of Monetary Economics*, forthcoming.

Brennan, M J and Schwartz, E S (1979), 'A continuous time approach to the pricing of bonds', *Journal of Banking and Finance*, Vol. 3, pages 133-56.

Canova, F and Ito, T (1991), 'The time-series properties of the risk premium in the Yen/Dollar exchange market', *Journal of Applied Econometrics*, Vol. 6, pages 125-42.

Cheung, Y W (1993), 'Exchange rate risk premiums', *Journal of International Money and Finance*, Vol. 12, pages 182-94.

Cox, J, Ingersoll, J and Ross, S (1985), 'A theory of the term structure of interest rates', *Econometrica*, pages 385-407.

Dai, Q and Singleton, K J (2000), 'Specification analysis of affine term structure models', *Journal of Finance*, Vol. LV, pages 1,943-78.

De Jong, F (2000), 'Time series and cross-section information in affine term structure models', *Journal of Business and Economic Statistics*, Vol. 18, pages 300-14.

De Jong, F and Santa-Clara, P (1999), 'The dynamics of the forward interest rate curve: a formulation with state variables', *Journal of Financial and Quantitative Analysis*, Vol. 34, pages 131-57.

Domowitz, I and Hakkio, C (1985), 'Conditional variance and the risk premium in the foreign exchange market', *Journal of International Economics*, Vol. 19, pages 47-66.

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

Engel, C (1996), 'The forward discount anomaly and the risk premium: a survey of recent research', *Journal of Empirical Finance*, Vol. 3, pages 123-92.

Fama, E (1984), 'Forward and spot exchange rates', *Journal of Monetary Economics*, Vol. 14, pages 319-38.

Groen, J J J and Balakrishnan, R (2005), 'Asset price based estimates of sterling exchange rate risk premia', *Bank of England Working Paper no. 250.*

Hai, W, Mark, N C and Wu, Y (1997), 'Understanding spot and forward exchange rate regressions', *Journal of Applied Econometrics*, Vol. 12, pages 715-34.

Hamilton, J D (1985), 'A standard error for the estimated state vector of a state-space model', *Journal of Econometrics*, Vol. 33, pages 387–97.

Hamilton, J D (1994), *Time series analysis*, Princeton University Press.

Harvey, A, Ruiz, E and Sentana, E (1992), 'Unobserved component time series models with ARCH disturbances', *Journal of Econometrics*, Vol. 52, pages 129-57.

Hodrick, R J (1989), 'U.S. international capital flows: perspectives from rational maximising models', *Carnegie-Rochester Conference Series on Public Policy*, Vol. 30, pages 231-88.

Hodrick, R J and Vassalou, M (2002), 'Do we need multi-country models to explain exchange rate and interest rate and bond return dynamics?', *Journal of Economic Dynamics and Control*, forthcoming.

Kaminsky, G and Peruga, R (1990), 'Can a time-varying risk premium explain excess returns in the forward market for foreign exchange?', *Journal of International Economics*, Vol. 28, pages 47-70.

Kim, C J and Nelson, C (2000), *State-space models with regime switching*, Cambridge, Mass., The MIT Press.

Leippold, M and Wu, L (2002), 'Asset pricing under the quadratic class', *Journal of Financial and Quantitative Analysis*, Vol. 5, pages 271-295.

Leippold, M and Wu, L (2003), 'Estimation and design of quadratic term structure models', *Review of Finance*, Vol. 7, pages 47-73.

Lewis, K L (1994), 'Puzzles in international financial markets', in Grossman, G and Rogoff, K (eds), *Handbook of international economics*, Amsterdam, North Holland.

Longstaff, F A and Schwartz, E S (1992), 'Interest rates volatility and the term structure: a two-factor general equilibrium model', *Journal of Finance*, Vol. 47, pages 1,259-82.

Lucas, R E, Jr (1982), 'Interest rates and currency prices in a two-country world', *Journal of Monetary Economics*, Vol. 10, pages 335-59.

Mark, N (1985), 'On time-varying risk premia in the foreign exchange market', *Journal of Monetary Economics*, Vol. 16, pages 3-18.

Modjtahedi, B (1991), 'Multiple maturities and time-varying risk premia in forward exchange markets: an econometric analysis', *Journal of International Economics*, Vol. 30, pages 69-86.

Newey, W and West, K D (1987), 'A simple positive-semi-definite heteroscedasticity and autocorrelation consistent covariance matrix', *Econometrica*, Vol. 55, pages 703-8.

Pearson, N D and Sun, T S (1994), 'Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll, and Ross model', *Journal of Finance*, Vol. 49, pages 1,279-304.

Taylor, M P (1988), 'A DYMIMIC model of forward foreign exchange risk, with estimates for three major exchange rates', *The Manchester School*, Vol. 56, pages 55-68.

Vasicek, O (1977), 'An equilibrium characterisation of the term structure', *Journal of Financial Economics*, Vol. 5, pages 177-88.

Watson, M and Engle, R (1983), 'Alternative algorithms for estimation of dynamic MIMIC, factor, and time varying coefficient regression models', *Journal of Econometrics*, Vol. 23, pages 385-400.