A term structure model of interest rates and forward premia: an alternative monetary approach
Daal, W.H.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 6

The Empirical Analysis

Our empirical analysis focuses on two objectives. First, we validate the structure of our general equilibrium model. The equilibrium conditions derived in previous chapter yield a number of testable propositions. In this respect we investigate whether the stochastic processes derived from our theoretical model provide an adequate description of the time series properties, i.e. the first and second moments, of the nominal interest rates and exchange rate. In addition, we test the cross-sectional restriction that our theoretical model imposes on the term structure of nominal interest rate. Second, our empirical analysis shows that our model provides an adequate description of the properties of the term structure of forward premia and accounts for the Fama condition.

To derive testable hypotheses we need to determine the exact conditional probability density function of the processes underlying these variables. For most of the variables in question the conditional densities are characterized by non-central chi-squared distributions. The disadvantage of these distribution functions is, as shown by Schroder (1989), that approximating algorithms have to be applied and that the conditional densities depend on the unobservable state variables in the model. We solve the problem of the unobservable state variables in two ways. One approach is to invert the three state variables, as in Longstaff and Schwartz (1992), from the set of equations (5.4), (5.6), and (5.7), which form a simple system of three linear equations in the three state
variables. Solving this set of linear equations we obtain explicit expression for the three unobservable state variables [see Equations (0.84), (0.82), and (0.83) in Appendix B.3],

\[
Y(t) = -c_1(R_h(t) - R_f(t) - (\mu_{m_h}^* - \mu_{m_f}^*)) + c_2 V_{dR_h}(t) - c_3 V_{dR_f}(t), \tag{6.1}
\]

\[
X_h(t) = c_4(R_h(t) - R_f(t) - (\mu_{m_h}^* - \mu_{m_f}^*)) + c_5 V_{dR_h}(t) + c_6 V_{dR_f}(t), \tag{6.2}
\]

\[
X_f(t) = c_7(R_h(t) - R_f(t) - (\mu_{m_h}^* - \mu_{m_f}^*)) - c_8 V_{dR_h}(t) + c_9 V_{dR_f}(t), \tag{6.3}
\]

where

\[
c_1 = \gamma_h^2 \sigma_{x_h}^2 \sigma_{x_f}^2 / \Psi^*
\]

\[
c_2 = \gamma_h^2 \sigma_{x_h}^2 \sigma_{x_f}^2 / \Psi^*
\]

\[
c_3 = \gamma_f^2 \gamma_h^2 \sigma_{x_h}^2 / \Psi^*
\]

\[
c_4 = \gamma_f^2 \sigma_{x_f}^2 \lambda_h^2 \sigma_{y}^2 / \Psi^*
\]

\[
c_5 = \gamma_f^2 \left( \lambda_f^2 \sigma_{y}^2 - \gamma_f^2 \sigma_{x_f}^2 \left( \lambda_f^2 - \lambda_h^2 \right) \right) / \Psi^*
\]

\[
c_6 = \gamma_f^2 \lambda_h^2 \sigma_{y}^2 / \Psi^*
\]

\[
c_7 = \gamma_h^2 \sigma_{x_h}^2 \lambda_f^2 \sigma_{y}^2 / \Psi^*
\]

\[
c_8 = \left( \gamma_h^2 \lambda_f^2 \sigma_{y}^2 + \gamma_h^2 \sigma_{x_h}^2 \left( \lambda_f^2 - \lambda_h^2 \right) \right) / \Psi^*
\]

\[
\Psi^* = \gamma_h^2 \gamma_f^2 \left( -\gamma_h^2 \sigma_{x_h}^2 \lambda_f^2 \sigma_{y}^2 + \gamma_f^2 \sigma_{x_f}^2 \left( \lambda_f^2 \sigma_{y}^2 + \gamma_h^2 \sigma_{x_h}^2 \left( \lambda_f^2 - \lambda_h^2 \right) \right) \right)
\]

This inversion of the state variables holds provided \( \Psi^* \neq 0 \), and therefore we can make a change of variable from \( Y(t) \) and \( X_i(t) \) to the nominal interest rate differential, \( R_h(t) - R_f(t) \), and the nominal interest rates volatilities, \( V_{dR_h}(t) \) and \( V_{dR_f}(t) \). This set of reduced form equations for the state variables introduces a complex non-linearity in the model parameters. We use this solution methodology for the validation of the structure of our model in Section 6.3. Note that the state variables in the reduced form equations depend on the unobserved volatilities of changes in the nominal interest rates. In this context, we use in Section 6.2 the GARCH framework, introduced by Engle (1982) and Bollerslev (1986), to estimate these unobserved volatilities. To account empirically for
the Fama conditions (Section 6.4), the model parameters have to be determined. In this case we use a second approach to solve for the unobservable state variables. We use the unconditional moments of these state variables. Since these state variables are stationary, this approach reduces the degree of non-linearity in the model parameters.

6.1 The data

The dynamic and cross-sectional restrictions of our cross-country term structure model are tested with monthly data for the United States and United Kingdom. We use a number of maturity yields drawn from the London Euro-currency market. In particular, our sample consists of one-month, 6-month, and one-year London Euro-currency interest rates for these countries over the sample period 1981 - 2001. The short-term riskless rate of interest for each of these countries are proxied by their respective one month interest rate. The currency used in this empirical study is the British pound, which is quoted as the dollar price of the foreign currency. We estimate the volatility of the nominal riskless interest rate by using the data for the one month Euro-currency rate. The observations of Euro-dollar, Euro-pound rates, and the exchange rate were obtained from the Datastream. The exchange rate was quoted by Bankers Trust and interest rates are middle quotes for Euro-currency deposits at the close of the London market.

6.1.1 Time series properties of the data

Table 1 presents the descriptive statistics for the variables used in this empirical analysis as well as the well-known regression results underlying the currency prices. The unconditional average level of the one-month, six-month, and one year US-yields are 7.21%, 7.44%, and 7.64%, respectively, expressed in annualized form. The same statistics for the UK-yields are on average 9.21%, 9.24%, and 9.31%, respectively, on annual basis. The unconditional mean of the rate of depreciation of the currency in question is consistent with the well-known empirical result [see, e.g., surveys by Hodrick (1987), Canova and
Marrinan (1995), and Engel (1996)], that is the rate of depreciation have means that are indistinguishable from zero. A more striking feature is the very high variability of this series, which exhibit standard deviation of the order of 15 times its mean. The series displays little persistence, i.e. auto-correlation.

In contrast, interest rates and interest rate differentials are less volatile - both in absolute terms and relative to their means - but are highly autocorrelated. The standard deviations of the US- and UK-yields are on average 3.3% and 3.1%, respectively. The nominal short-term interest rates of the countries are highly correlated, and the nominal yields of the longer maturities are more strongly correlated.

A number of empirical studies in the economic and finance literature have found that high-frequency changes in the exchange rates are described by distributions with fatter tails than the normal distribution [see for example Westerfield (1977)]. Saá-Raquejo (1994) put forth two alternative explanations that account for the observed leptokurtosis: (1) that the changes in the logarithms of exchange rates are independently identically distributed drawn from a non-normal fat tailed distribution, or (2) that the changes are generated by distributions whose parameters vary over time.

A glance at the graph of the change in the spot exchange rate clearly shows tranquil periods and periods of volatile behavior (see figure 3). In order to investigate this, autocorrelations are computed for the squared-valued series. Under the null hypothesis of conditional homoskedasticity these autocorrelations should equal zero. This is strongly rejected by the data, suggesting the presence of time-varying conditional variances. We also find evidence on the presence of leptokurtosis and heteroskedasticity in the series. A formal statistical test for this particular pattern of time variation in volatility is the Lagrange Multiplier (LM) test on ARCH(p). The test results indicate the presence of ARCH effects in the dynamics of the spot exchange rate (see Table 1). This is also confirmed by the presence of serial correlation of the first order in the squared residuals. Furthermore, we clearly find proof of a heavy tailed distribution (leptokurtosis).

The fourth moment of the UK-pound equals 4.9, which is larger than the value of 3,
the kurtosis of a normal distribution. This means that the tail of the distribution of this exchange rate is fatter than the tail of a normal distribution, which basically implies that large observations occur more often than normally expected. In addition it should be noted that the series exhibit a small negative skewness, which implies that the left tail of the distribution is slightly fatter than the right tail or, stated differently, large depreciation tends to occur slightly more often than large appreciations. Heteroskedasticity and leptokurtosis are stronger for the forward premium than they are for the depreciation rate of the currency. This result is consistent with what is reported by Bekaert (1995) and gives evidence of time varying risk premia. The US-interest rates also displays evidence of the presence of leptokurtosis. This can be ascribed to the increased idiosyncratic variation in the US-yields since the early 1980s.

Before testing our model it is necessary to establish whether the variables in our model are integrated or co-integrated in the sense of Granger (1981) and Engle and Granger (1987). The importance of non-stationarity is that it may affect the estimation procedure, like the correct specification of the testable hypothesis and statistical tests to use. Estimating models with integrated variables may lead to non-stationary errors, and thus to invalid asymptotic inferences. The alternative of making the variables stationary by taking differences may lead to lost of information, such as important long term relations. We apply the testing framework developed by Phillips (1987), Phillips and Perron (1988), and Perron (1988,) as it allow us to control for autocorrelation of the unknown form and heteroskedasticity. Their framework allow for the control of serial correlation in the series, by adjusting the test statistics to account for the serial correlation in the innovations of the processes under study. This is the main advantage of their framework compared with the augmented Dickey and Fuller (1979) tests, which corrects for serial correlation by adding differenced terms as explanatory variables in the regression.

Table 1 presents the results of the Phillips and Perron unit roots tests for the short-term riskless interest rates, the six-month and one-year yields for the different countries in the sample, the rate of depreciation of the spot exchange of the UK-pound against
the US-dollar, and the forward premium. All series except the forward premium and the rate of depreciation of the spot exchange rate appear to be integrated and therefore non-stationary. The interest rate series are integrated of order one, \( I(1) \). Therefore, all cross-country pairs on interest rates with the same maturity are cointegrated. Hence, as argued by Raquejo (1994), when the interest rates for both countries (e.g. the interest rate differential) are included as explanatory variables in an equation, where the dependent variable is stationary, their influence will enter through their co-integrating relation, which is stationary by definition. Since the spot and forward exchange rates are integrated of order one, that is non-stationary with a unit root, both the rate of depreciation of the spot exchange rate and the forward foreign exchange premium are stationary [as in Baillie and Bollerslev (1989)]. Therefore we can state that spot and forward exchange rates are co-integrated of order one. Based on these results we know that the standard asymptotic inference procedures for our model are valid.

Finally, and most importantly for our purposes, the regression results for the currency price is in accordance with existent empirical studies of exchange rate [see Engel (1996)]. Table 1 shows that the data are supportive of the covered interest rate parity (CIP) condition, since this would otherwise imply the existence of riskless arbitrage opportunities. Similar results were obtained by Fratianni and Wakeman (1982). In addition we note that, in conformity with general empirical findings, the uncovered interest rate parity (UIP) is not supported by the data, that is the estimates of the Fama slope coefficient is not only unequal to one but also negative. Panel II of Table 1 shows the regression results for the UIP. The estimated value of the slope coefficient from the forward premium regression is significantly different from one, namely -2.726 for the UK-pound.

### 6.2 The interest rate volatilities

To obtain estimates of the unobservable volatility of the spot interest rates several methodologies have been used in literature, such as the 'historical volatility' or the 'im-
plied volatility'. The first approach uses historical data of interest rates to calculate the
volatility of the interest rate. The disadvantage of this approach is that it does not pro-
vide current information on the interest rate volatility. The implied volatility approach
means that the actual bond prices are set equal to their theoretical values. The implied
volatility is obtained by solving the resulting system of non-linear equations. As argued
by Longstaff and Schwartz (1992) this approach depends on the functional form of the
bond prices implied by our model and it is not suitable for our objective, i.e. validating
the properties of our model. Therefore, to avoid possible biasing of the results in favor
of our model we use an approach that is independent of our model to obtain estimates
of the interest rate volatility.

In line with Longstaff and Schwartz (1992), Ball and Torous (1999), Chan et al.
(1992), and Koedijk et al. (1997) we use a GARCH framework introduced by Engle
(1982) and Bollerslev (1986) to estimate the volatility of the nominal interest rates,
\( V_{dR_b} (t) \) and \( V_{dR_f} (t) \). It is well-known that these models are capable of capturing many
of the properties of financial time series, such as volatility clustering and excess kurtosis.
Consider an observed time series \( y_t \), which can be written as the sum of a predictable
part and a stochastic part,

\[
y_t = E[y_t | \Omega_{t-1}] + \varepsilon_t, \tag{6.4}
\]

with \( \Omega_{t-1} \) denoting the relevant information set available at time \( t - 1 \). In this section
we will be considering the case where the conditional variance of \( \varepsilon_t \) is time-varying, that
is \( E[\varepsilon_t^2 | \Omega_{t-1}] = V_t \). This condition of heteroskedasticity can be formulated as

\[
\varepsilon_t = z_t \sqrt{V_t}, \tag{6.5}
\]

where \( z_t \sim I.I.D (0,1) \) and \( V_t \) is a non-negative function of \( \Omega_{t-1} \). Next, we consider
two functional forms for this conditional variance, \( V_t \), namely a linear and a non-linear
specification.

Engle (1982) formulated a basic linear auto-regressive conditional heteroskedastic
(ARCH) function of $V_t$ to describe the conditional variance of the shocks that occurs at time $t$

$$V_t = \beta_1 + \beta_2 \varepsilon_{t-1}^2,$$  \hspace{1cm} (6.6)

where, given the non-negativity condition imposed on $V_t$, the parameters has to satisfy the following requirements: $\beta_1 > 0$ and $\beta_2 \geq 0$. If $\beta_2 = 0$, then the conditional variance is constant and the time series is homoskedastic. It can be noted from this specification, for $\beta_2 > 0$, that a large shock in one period will be followed by another large shock in the next period. It is this feature that allows this model to capture the observed volatility clustering in financial data. By taking the fourth moment of $\varepsilon_t$, it can be shown that this model is able to capture that excess kurtosis present in financial series. Note, however, that this specification due to the quadratic specification, is independent of the sign of the shock.

Bollerslev (1986) proposed a further generalization that can capture the dynamic patterns in the conditional volatility without having to recur to the cumbersome lag structure of an ARCH($q$) model. This so-called Generalized ARCH (GARCH) model of order (1,1) can be written as

$$V_t = \beta_1 + \beta_2 \varepsilon_{t-1}^2 + \beta_3 V_{t-1}$$  \hspace{1cm} (6.7)

whereby the condition of non-negativity of $h_t$ requires the following parameter restrictions: $\beta_1 > 0$, $\beta_2 \geq 0$, and $\beta_3 \geq 0$. Furthermore, it can be shown that for $\beta_3$ to be identified it is necessary for $\beta_2 > 0$.\footnote{As showed by Franses and van Dijk (1999), if the sum of $\beta_2$ and $\beta_3$ is close or equal to one, what is very common in high-frequency financial time series, it implies the presence of unit root in the conditional variance and in this case the model can be referred to as Integrated GARCH [IGARCH]. Furthermore, it is shown that, although the IGARCH(1,1) model is not covariance stationary, it may still be strictly stationary.}

Nelson (1991) argues that GARCH models applied on financial time series have in general three shortcomings. Firstly, these models by assumption exclude the possibility of correlation between large negative shocks and periods of high volatility, which suggests
that positive and negative shocks may have asymmetric impact on the conditional volatility in the next period. Stated differently, GARCH models can capture the magnitude of shocks, but they are unable to distinguished between the sign of the shocks. This is due to the fact that only the square of the shocks is incorporated in the GARCH model. This is in contrast with the evidence found in empirical research, namely that "bad news", meaning for example a depreciation of the exchange rate or lower excess return, leads to a volatility increase and that "good news" may lead to a decrease in volatility. Secondly, he stated that GARCH models impose parameter restrictions in order to guarantee the non-negativity condition of the conditional variance, that are often violated by estimated coefficients. Additionally these restrictions may lead to an incorrect representation of the dynamics of the conditional variance process. Thirdly, he argued that it is difficult to interpret whether shocks to the conditional variance are persistent or not. Based on these arguments a new class of GARCH model was proposed by Nelson (1991),

\[
\ln (V_t) = \beta_1 + \beta_2 (|z_{t-1}| - E (|z_{t-1}|)) + \beta_3 z_{t-1} + \beta_4 \ln (V_{t-1}).
\]  

(6.8)

This exponential GARCH [EGARCH] model is capable of capturing the asymmetric volatility shocks present in most financial time series. First of all it should be noted that this asymmetric effect runs through the \( \beta_3 \) parameter. If \( \beta_3 = 0 \) this model is reduced to a symmetric model as GARCH, given that positive shocks have the same impact as negative shocks on the conditional variance. Positive shocks have a smaller effect on volatility than negative shocks if \(-1 < \beta_3 < 0\) and if \(\beta_3 < -1\) positive shocks will reduce volatility while negative shocks will increase volatility. The fact that \(\beta_3\) can capture this type of effects is called the leverage effect. Another advantage of Nelson's specification is that, since it describes the relation between the logarithm of the conditional variance and past shocks, no restriction need to be imposed on the parameter specifications in order to guarantee the non-negativity condition of \(V_t\).

To estimate the nominal interest rate volatilities we model, as in Longstaff and Schwartz (1993), discrete changes in the riskless spot interest rate by the following econo-
metric specification

\[ R_{i,t+1} - R_{i,t} = a_0 + a_1 R_{i,t} + a_2 V_{i,t} + \varepsilon_{i,t+1}, \quad (6.9) \]

where \( \varepsilon_{i,t+1} = z_{i,t+1} \sqrt{V_{i,t+1}} \) is the prediction error of the riskless spot interest rate \( i \) and \( \varepsilon_{i,t} \sim N(0, V_{i,t}) \). The error term, \( z_{i,t+1} \), is normalized to have variance equal to one. This discrete time specification is only an approach to the continuous time formulation in Equation (5.5) and, therefore, the GARCH parameter estimates need not map directly to the parameters of the continuous-time process. However, there is a functional equivalence, as the discrete-time approach implies that changes in the nominal interest rates is a function of the interest rate level and its volatility.

We consider two types of GARCH models, representing the linear and non-linear specification, to estimate the conditional variance of the spot interest rates. An extended Bollerslev (1991) symmetric specification of a nonnegative function for \( V_{i,t} \) [Equation (6.7)], as applied in Longstaff and Schwartz (1992),

\[ V_{i,t} = b_0 + b_1 R_{i,t-1} + b_2 V_{t-1} + b_3 \varepsilon_{t-1}^2, \quad (6.10) \]

where \( b_0 > 0, b_1 \geq 0, \) and \( b_2 > 0 \) and an extended Nelson's (1991) asymmetric representation [Equation (6.8)],\(^2\)

\[ \ln (V_{i,t}) = b_0 + b_1 R_{i,t-1} + b_2 \ln (V_{i,t-1}) + b_3 (|z_{t-1}|) + b_4 z_{t-1}. \quad (6.11) \]

Both specifications extends the standard GARCH-M and EGARCH formulation, respectively, by incorporating the level of the riskless spot interest rate in the volatility equation. The results of both specification are presented in Table 2 below.

In line with Longstaff and Schwartz (1992) we used the BHHH, that is the Berndt-

\(^2\)This is a slightly adjusted specification compared to equation (6.8), which under the assumption of normal errors will yield identical parameter estimates, with exception of the intercept.
Hall-Hall-Hausman (1974), numerical algorithm to find the quasi maximum likelihood parameter estimate for the GARCH specification. To ensure that the estimates of the interest rate volatilities were robust and that the algorithm converged to the global maximum we also used different starting values. Table 2 contains the results of the GARCH (1,1) specification for our data series. The results reported for the GARCH(1,1) model show that all parameter estimates are significant, even after applying a robust standard error procedure. According to the Lagrange Multiplier (LM) test for the presence of ARCH effects in the dynamics of the interest rates the residuals do not contain any ARCH element. Furthermore, the results reveal the typical findings in empirical applications of GARCH specification, namely that the estimate of $b_2$ is fairly small, the estimate of $b_3$ is rather large, and that the sum of these two, $b_2 + b_3$, is close to unity. This implies that the conditional volatility is persistent as the shocks to the conditional variance decay at a very slow pace. This model can be referred to as an Integrated GARCH (IGARCH) model. Despite the fact that the constant term, $b_0$, is significantly different from zero, it is not positive as required.

Before regressing the EGARCH specification a formal test for asymmetric GARCH is performed on the dynamics of the short term riskless interest rate series, that is the ‘joint sign and size test’ as proposed by Engle and Ng (1993). Following Franses and van Dijk (1999) we construct a dummy variable, $S_t^-$, that takes the value of 1 when $\hat{\varepsilon}_t$ is negative and zero otherwise. The test indicates whether the squared residual $\hat{\varepsilon}_t^2$ can be predicted by $S_{t-1}^{-}$, $S_{t-1}^{-}\hat{\varepsilon}_{t-1}$, and $S_{t-1}^{+}\hat{\varepsilon}_{t-1}$, where $S_{t-1}^{+} \equiv 1 - S_{t-1}^{-}$. These represent the Sign Bias (SB), the Negative Size Bias (NSB), and the Positive Size Bias (PSB), respectively. Thus the test is performed on the following equation

$$\hat{\varepsilon}_t^2 = \phi_0 + \phi_1 S_{t-1}^- + \phi_2 S_{t-1}^- \hat{\varepsilon}_{t-1} + \phi_3 S_{t-1}^+ \hat{\varepsilon}_{t-1} + \zeta_t,$$  \hspace{1cm} (6.12)

under the assumption of conditional homoskedasticity, that is with $H_0 : \phi_1 = \phi_2 = \phi_3 = 0$. Table 3 reports the estimation results for the joint sign and size test as applied on the dynamics of the short-term riskless US and UK interest rate. The joint-test results
indicate that the null hypothesis should be rejected for both the US and the UK interest rates with a $p$-value of 0.00, which confirms the presence of asymmetric GARCH in both series. Noteworthy, is that for both interest rates only the parameter estimate for PSB is significantly different from zero, which implies that there is a tendency that high volatility periods begin with a positive innovative shock in the interest rate processes.

As in the case of the linear GARCH specification we used the BHHH numerical algorithm to find the maximum likelihood parameter estimate for the EGARCH specification and different starting values. The parameter estimates for the EGARCH specification [Equation (6.11)] confirm the results of the sign test. The estimate for $b_4$ is positive and significant, which reflects the presence of leverage effect in the non-linear specification and that positive shocks have a larger effect on interest rate volatility than negative shocks. As can be seen from table 2, the residuals do not contain any ARCH element as indicated by the ARCH-LM test statistics. This is also confirmed by the Ljung-Box test statistics of 6.78, which indicates the absence of autocorrelation in the residual term.

The results of both estimates and the joint sign and size test indicates that probably the EGARCH estimation procedure provides a better description of the short term nominal interest rate volatility. In addition note that the conditions for strict stationarity and covariance stationarity of the EGARCH volatility process are identical, while the GARCH models allows for strictly stationary processes that are not covariance-stationary ($b_2+b_3$ is close to unity). This feature of the GARCH specification is not critical. It is, as noted by Andersen and Lund (1997), the additional interaction via the interest rate level effect and the near integration in the mean dynamics, which will induce further destabilizing elements into the interest rate dynamics. A more formal criterium for comparing the performance of these two models, the Akaike Information Criteria (AIC) and the Schwarz Information Criteria (SIC), indicates that the EGARCH(1,1) specification has a slightly better in-sample-fit compared with the GARCH(1,1) specification. Therefore, we use the EGARCH to obtain an estimate of the volatility of the short term riskless rate of interest of both countries. Figure 1 and 2 compares the estimated standard deviation of
changes in the short-term riskless nominal interest rates with the absolute changes in the respective interest rates. The estimated standard deviation is measured by the square root of the conditional variance, \( V_{t,t} \). Note that the absolute changes in the short-term riskless rate of interest rates is an ex-post measure of its volatility. The EGARCH interest rate volatility estimate provides a description of the ex-post volatility pattern of both series.

### 6.3 The econometric specification

The model developed in Chapter 3 imposes two types of restriction on the data generating process, namely the dynamic properties and the cross-sectional restrictions. In this section, we describe the econometric approach used in examining both conditions. We provide an econometric specification of the stochastic process derived from our model for the equilibrium nominal interest rates and the rate of depreciation. Ultimately the cross-sectional properties of our model are formulated as a set of testable restrictions.

#### 6.3.1 The estimation framework: GMM

Our econometric approach is to test the dynamic specification and the cross-sectional properties of our model as a set of over-identifying restrictions on a system of moment equations using the Generalized Method of Moments (GMM) approach of Hansen (1982). Several studies of the interest rate and exchange rate dynamics have applied the GMM approach [e.g., Gibbons and Ramaswamy (1993), Harvey (1988), Longstaff and Schwartz (1992), Chan et al. (1992), and Saá-Raquejo (1994)]. As Chan et. al. (1992) point out, Hansen’s (1982) GMM framework provides a number of important advantages, which make it an intuitive and logical choice for estimating these continuous-time processes. Since these advantages also hold for our framework, we summarize them as follows. First, the GMM approach does not require that changes in the nominal interest rates and the rate of depreciation of the spot exchange rate are normally distributed. The
asymptotic justification for the GMM procedure requires only that the distribution of the processes in question be stationary and ergodic and that the relevant expectations exist. This is of particular importance in testing of our continuous-time model, since the distributions of the underlying processes are characterized by non-central Chi-squared distributions, which are highly non-linear and require the use of approximating algorithms. Second, the GMM-estimators and their standard errors are consistent even if the disturbance term, are conditionally heteroskedastic and serially correlated. Since the temporal aggregation problem that arises from estimation of continuous-time processes with discrete-time data are likely to influence the distribution of the disturbance terms, the GMM approach should further alleviate the impact of this approximation error on the parameter estimates.

The GMM-procedure of Hansen (1982) can be summarized as follows. As in Hamilton (1994) consider a \((k \times 1)\) vector-valued function \(h(\theta, x_t)\), where \(x_t\) is a \((l \times 1)\) vector of explanatory variables, and \(\theta\) a \((m \times 1)\) vector of unknown parameters. Suppose that the value of \(\theta\), denoted by \(\theta_0\), can be characterized by the \(k\) moments conditions or orthogonality conditions

\[ E[h(\theta_0, x_t)] = 0. \]  

(6.13)

Then the GMM procedure consists of estimating the value of \(\theta\), i.e. \(\hat{\theta}_n\), that minimizes the quadratic criterion function,

\[ \hat{\theta}_n = \arg\min_{\theta} g(\theta; Y_n)' W_n g(\theta; Y_n). \]  

(6.14)

where \(g(\theta; Y_n)\) is the sample moments, given by

\[ g(\theta; Y_n) = \frac{1}{n} \sum_{t=1}^{n} h(\theta, x_t) \]

\(Y_n\) is an \((nl \times 1)\) vector of the observed sample, and \(W_n\) is a \((k \times k)\) symmetric positive definite weighting matrix. The parameter estimate \(\hat{\theta}_n\) can be determined as the solution
to the system of first order condition for the function $Q(\theta;\gamma_n)$ in equation (6.14). Matrix differentiation implies that minimization of $g(\theta;\gamma_n)$ with respect to $\theta$ is equivalent to solving the homogeneous system of orthogonality conditions,

$$G(\theta;\gamma_n)' W_n G(\theta;\gamma_n) = 0,$$

(6.15)

where $G(\theta;\gamma_n)$ is the $(k \times m)$ Jacobian matrix of $g(\theta;\gamma_n)$, i.e. the matrix containing the partial derivatives, with respect to $\theta$. Hansen (1982) shows that the GMM estimator of $\theta$ with the smallest asymptotic covariance matrix is obtained when the optimal weighting matrix $W_n$ is given by the inverse of the (asymptotic) covariance matrix of $g(\theta;\gamma_n)$. That is $W_n = S^{-1}$ with

$$S(\theta_0;\gamma_n) = \lim_{n \to \infty} n E \left[ g(\theta_0;\gamma_n)(\theta_0;\gamma_n)' \right]$$

(6.16)

If the moment conditions are serially uncorrelated, $S$ can be estimated as

$$\hat{S}_n(\hat{\theta}_n, x_t) = \frac{1}{n} \sum_{t=1}^{n} h(\hat{\theta}_n, x_t) h(\hat{\theta}_n, x_t)',$$

(6.17)

where $\hat{\theta}_n$ is a consistent estimate of $\theta_0$. We choose an estimate of the covariance matrix that is robust to heteroskedasticity and autocorrelation of unknown form. To compute the weighting matrix we use the Bartlett kernel to weight the autocovariances and the Newey-West fixed bandwidth. The minimized value of the quadratic criterion function in equation (6.14) is, as shown by Hansen (1982), asymptotically distributed as a $\chi^2(k - m)$ under the null hypothesis that the restrictions are true.

### 6.3.2 The dynamics of the nominal interest rates and the spot foreign exchange rate

The equilibrium conditions of our model imply that the continuous-time nominal interest rate processes and the rate of depreciation of the spot exchange rate in equations (5.5) and (5.19) depend explicitly on the level of the nominal interest rate differential and the
volatilities of both riskless spot interest rates. Therefore, as in Brennan and Schwartz (1982), Dietrich-Campbell and Schwartz (1986), Sanders and Unal (1988), and Chan, Karolyi, Longstaff, and Sanders (1992) we approximate the continuous-time nominal interest rate and exchange rate processes in equations (5.5) and (5.19) by, respectively, the following discrete-time econometric specification

\[ R_{h,t+1} - R_{h,t} = \beta_{11} + \beta_{12} R_{diff,t} + \beta_{13} V_{dR_{h,t}} + \beta_{14} V_{dR_{f,t}} + \varepsilon_{R_{h,t+1}} \]  
\( (6.18) \)

\[ R_{f,t+1} - R_{f,t} = \beta_{21} + \beta_{22} R_{diff,t} + \beta_{23} V_{dR_{h,t}} + \beta_{24} V_{dR_{f,t}} + \varepsilon_{R_{f,t+1}} \]  
\( (6.19) \)

\[ e_{t+1} - e_t = \beta_{31} + \beta_{32} R_{diff,t} + \beta_{33} V_{dR_{h,t}} + \beta_{34} V_{dR_{f,t}} + \varepsilon_{e,t+1}, \]  
\( (6.20) \)

where \( R_{diff,t} = R_{h,t} - R_{f,t} \), \( e_t \) is the log spot foreign exchange rate and

\[ E[\varepsilon_{R_{h,t+1}}] = 0, \quad E[\varepsilon_{R_{h,t+1}}^2] = \xi_{11} + \xi_{12} R_{diff,t} + \xi_{13} V_{dR_{h,t}} + \xi_{14} V_{dR_{f,t}}, \]

\[ E[\varepsilon_{R_{f,t+1}}] = 0, \quad E[\varepsilon_{R_{f,t+1}}^2] = \xi_{21} + \xi_{22} R_{diff,t} + \xi_{23} V_{dR_{h,t}} + \xi_{24} V_{dR_{f,t}}, \]

\[ E[\varepsilon_{e,t+1}] = 0, \quad E[\varepsilon_{e,t+1}^2] = \xi_{31} + \xi_{32} R_{diff,t} + \xi_{33} V_{dR_{h,t}} + \xi_{34} V_{dR_{f,t}}. \]

This econometric specification corresponds with the dynamics of equilibrium nominal interest rates and the depreciation rate. The linear drift components and the variances of the changes in the nominal interest rates and the rate of depreciation are allowed to depend on the interest rate differential and the interest rate volatilities in a way that is consistent with the continuous-time equilibrium model. Note, however, that the discretized processes in equations (6.18), (6.19) and (6.20) are only approximations of the continuous-time specification. As noted by Chan et al. (1992) this discretization error is of second-order importance if changes in the level of the nominal interest rates and the spot exchange rate are measured over short periods of time. To validate the dynamic properties of our model we test the first and second moment of the interest rate and exchange rate processes. The parameter vectors \( \beta = [\beta_{ij}] \) and \( \xi = [\xi_{ij}] \), for
\( i = [1, 2, 3] \) and \( j = [1, \ldots, 4] \), describe the linear drift components and the variances of these processes. The \( i = [1, 2, 3] \) denote, respectively, the domestic nominal interest rate, the foreign nominal interest rate, and the exchange rate.

Define \( y_t \) to be a \((3 \times 1)\) vector consisting of \( y_{i,t} \) for \( i = R_{h,t}, R_{f,t}, \) and \( e_t \). Let \( e_t \) denote a vector of residuals, with its elements representing the deviation of the observed value of \( \Delta y_t \) from the theoretical value implied by equations (5.5) and (5.19). The elements of \( e_t \) are defined as

\[
e_{i,(t+1)} = y_{i,(t+1)} - y_{i,t} - (\beta_{i1} + \beta_{i2} R_{diff,t} + \beta_{i3} V_{dR,t} + \beta_{i4} V_{dR,t}) .
\]

We define \( \theta_Y \) to be the parameter vector with elements \( \beta_i \) and \( \xi_i \) and let the vector \( h_t(\theta_Y) \) be

\[
h_t(\theta_Y) = \begin{bmatrix}
\varepsilon_{i,(t+1)} \\
\varepsilon_{i,(t+1)} R_{diff,t} \\
\varepsilon_{i,(t+1)} V_{dR,t} \\
\varepsilon_{i,(t+1)} V_{dR,t} \\
\varepsilon_{i,(t+1)}^2 - \xi_i x_t' \\
(\varepsilon_{i,(t+1)}^2 - \xi_i x_t') R_{diff,t} \\
(\varepsilon_{i,(t+1)}^2 - \xi_i x_t') V_{dR,t} \\
(\varepsilon_{i,(t+1)}^2 - \xi_i x_t') V_{dR,t}
\end{bmatrix},
\]

where

\[
z_t = (1, R_{diff,t}, V_{dR,t}, V_{dR,t}) .
\]

Let \( f_t(\theta_Y) \) be a \((8m \times 1)\) vector formed by stacking the \( h_t(\theta_Y) \) vectors for the \( m \) different variables, these are the domestic nominal interest rate, the foreign nominal interest rate, and the rate of depreciation of the spot exchange rate. The theoretical model developed in this study implies that \( E[f_t(\theta_Y)] = 0 \). This equation represents a set of 24 population orthogonality conditions from which an estimator of the parameter
vector $\hat{\theta}_Y$ can be obtained through the GMM procedure. In order to construct a test of over-identifying restrictions implied by our model, we include the lagged explanatory variables in the vector of instruments. The test of over-identifying restrictions can be interpreted as testing whether the three reduced-form state variables, $R_{dtf,t}, V_{dR_h,t},$ and $V_{dR_R,t}$, not only have explanatory power for the nominal interest rates and the exchange rate dynamics, but that they have it in the way predicted by our model.

As described above the GMM procedure consists of replacing $E[f_t(\theta_Y)] = 0$ by its sample counterpart, $g_n(\theta_Y)$, using the $n$ observations,

$$g_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} f_t(\theta_Y, x_i).$$

The minimized value of the quadratic form in equation (6.14) is distributed as a $\chi^2$ variate under the null hypothesis that the model is true with degrees of freedom equal to the number of orthogonality conditions net of the number of parameters to be estimated, that is with 18 degrees of freedom. This measure proved a goodness-of-fit test for the model. A high value for the test statistic means that the dynamic properties of the two-country term structure model are rejected (i.e. misspecified).

6.3.3 The specification of the term structure model of nominal interest rates

To perform the cross-sectional test on our term structure model of nominal interest rates we rewrite the expression for the equilibrium nominal yield in equation (5.16). Let $K_t$ be the yield on a $\tau$-maturity bond denominated in the $i$-th currency, than by using the expression for the state variables in equations (6.1), (6.2), and (6.3), we can rewrite the nominal term structure model for country $i$,
as a linear function of the nominal interest rates differential and the volatilities of both interest rates,

\[ K_{hr} = a_{h,r} R_{diff,t} + b_{h,r} V_{dR_h(t)} + c_{h,r} V_{dR_f(t)} + d_{h,r} \]

(6.22)

and

\[ K_{fr} = a_{f,r} R_{diff,t} + b_{f,r} V_{dR_h(t)} + c_{f,r} V_{dR_f(t)} + d_{f,r}, \]

(6.23)

where \( a_{i,r}, \ldots, d_{i,r} \) for \( i = h, f \), are maturity specific constants. Note that for \( \tau \to 0 \), \( K_{i,r} \to R_{i,t} \). The cross-country term structure of nominal interest rates developed in this study implies that we can model discrete changes in the observed yield values, \( K_{i,r} \), as linear functions of the changes in the nominal interest rate differential and the volatilities of both spot interest rates. Therefore we can, as in Longstaff and Schwartz (1992), express changes in observed values of yields as linear function of the changes in the redefined state variables,

\[ \Delta K_{ir} = a_{i,r} \Delta R_{diff,t} + b_{i,r} \Delta V_{dR_h(t)} + c_{i,r} \Delta V_{dR_f(t)} \text{ for } i = h, f, \]

(6.24)

where \( \Delta d_{i,r} \) is zero, since it is a maturity specific constant.

Let \( \varepsilon_{K_{i,r}} \) denote the deviation of the observed value of \( \Delta K_{i,r} \) from the theoretical value implied by equation (6.24),

\[ \varepsilon_{K_{i,r}} = \Delta K_{i,r} - a_{i,r} \Delta R_{diff,t} - b_{i,r} \Delta V_{dR_h(t)} - c_{i,r} \Delta V_{dR_f(t)}. \]

We define \( \theta_{K_{i,r}} \) to be the parameter vector with elements, \( a_{i,r} \), \( b_{i,r} \), and \( c_{i,r} \). Define the
vector $h_t(\theta_{K,t})$ as

$$h_t(\theta_{K,t}) = \begin{bmatrix} \varepsilon_{K,t} \\ \varepsilon_{K,t} \Delta R_{diff,t} \\ \varepsilon_{K,t} \Delta V_{Rn,t} \\ \varepsilon_{K,t} \Delta V_{Rf,t} \end{bmatrix}$$

and $f_t(\theta_K)$ as a $(4m \times 1)$ vector formed by stacking the $h_t(\theta_{K,t})$ vectors for the $2m$ different maturities, that is $m$ maturities for each country. The theoretical model developed in this study implies that $E[f_t(\theta_K)] = 0$. This equation represents a set of 24 population orthogonality conditions from which an estimator of the parameter vector $\hat{\theta}_K$ can be obtained through the GMM procedure. As described above the GMM procedure consists of replacing $E[f_t(\theta_K)] = 0$ by its sample counterpart, $g_n(\theta_K)$, using the $n$ observations,

$$g_n(\theta_K) = \frac{1}{n} \sum_{i=1}^{n} f_t(\theta_K, x_t).$$

The minimized value of the quadratic form in equation (6.14) is distributed as a $\chi^2$ variate under the null hypothesis that the model is true. For the term structure data we use the one month, 6 month, and 1 year continuously compounded interest rates (expressed on annual basis) for both the US and UK. The estimated volatility series is obtained from the GARCH-framework [see Section 4.2].

6.3.4 The term structure of foreign exchange returns: An empirical framework

In this sub-section we specify a framework that allows us to examine the conditions under which our equilibrium model of the term structure of foreign exchange returns is consistent with the data. That is, it allows us to determine the parameters values that account for the currency puzzle and simultaneously characterize the dynamic properties of the main equilibrium results of our model, i.e. the nominal interest rates and the rate of depreciation. Due to the complexity of the model (i.e. highly non-linear and many
model parameters) and the distributional properties of the underlying state variables it is quite difficult to estimate all the parameters of the model. As a result we proceed as follows in order to determine the model parameters that account for the currency puzzle.

Let $\Theta = \left[ \mu_R, \gamma^*_i, \lambda^*_i, \kappa^*_x, \kappa^*_y, \theta_x, \theta_y, \sigma_z, \sigma_y, \sigma_{\eta}, \alpha_z, \alpha_y, \alpha_x \right]$, for $i = f, h$, represent a set that contains all the parameters of the model. To reduce the parameter space we assume that $\sigma_{\eta} = \sigma_{x}$, and $(\sigma_{\eta, h} - \sigma_{\eta, f}) = \sigma_{\eta}$. The components of the terms $\alpha_y = \alpha_{y_h} - \alpha_{y_f} - (\sigma_{\eta, h} - \sigma_{\eta, f}) \sigma_{\eta, f}$ and $\mu_R = \rho + \mu_{m_i} - \sigma^2_{m_i}$ cannot be identified separately in this framework, therefore we determine this term as an aggregate. Note that this does not alter the results of our model, since it is the sign and the magnitude of these parameters that are relevant in accounting for the currency puzzle. As mentioned in Chapter II, most equilibrium models of the forward premium face a trade-off when explaining the currency puzzle. That is, they either allow for implausible parameter values, which ultimately implies, for example, a positive probability of negative nominal interest rates, or that their model cannot generate a large volatility of the forward premia to account for the puzzle. Our methodology is therefore aimed at calibrating the model parameters such that they account simultaneously for the currency puzzle and the time series properties of the underlying variables, that is the unconditional mean, variances, and co-variances. Therefore, we determine the parameter values of $\Theta$ by using the estimated slope coefficient from the regression of the term structure of foreign exchange returns on the yield differentials; the unconditional mean of the nominal interest rates, the instantaneous rate of depreciation and of the volatilities of the changes in the nominal interest rates and the spot currency price; the unconditional co-variances between these variables; and the unconditional variances of the nominal interest rates.

To examine the implications for the forward premia note that we can write the $\tau$-
period foreign exchange returns as

\[ e_T - e_t = \beta_{1T} + \beta_{2T} (f_{eT} - e_t) + \epsilon_{eT} \]  

(6.25)

where \( \tau = T - t \) and the parameters \( \beta_{1T} \) and \( \beta_{2T} \) are maturity specific constants. From a theoretical perspective one expects \( \beta_{1T} = 0 \) and \( \beta_{2T} = 1 \), which implies that \( f_{eT} = E_t(e_T) \).

From Equation (5.29) we know that the equilibrium term structure of forward premia in log-specification is defined as

\[ f_{eT} - e_t = \log P_{fr}(R_f) - \log P_{hr}(R_h) = K_{hr} - K_{fr}, \]  

(6.26)

By substituting equation (6.26) in the equation for the \( \tau \)-period foreign exchange rate returns we obtain

\[ e_T - e_t = \beta_{1T} + \beta_{2T} (K_{hr} - K_{fr}) + \epsilon_{eT}. \]  

(6.27)

Define a vector of parameters, \( \theta_{eT} \), with elements \( \beta_{1T} \) and \( \beta_{2T} \). Let the vector \( h_t(\theta) \) be

\[ h_t(\theta_{eT}) = \begin{bmatrix} \epsilon_{eT} \\ \epsilon_{eT} (K_{hr} - K_{fr}) \end{bmatrix}, \]

where the deviations of the exchange rate at time \( T \) from the forward foreign exchange rate determined at time \( t \) for delivery at time \( T \) can be written as

\[ \epsilon_{eT} = e_T - e_t - \beta_{1T} - \beta_{2T} (K_{hr} - K_{fr}). \]

By stacking the \( h_t(\theta_{eT}) \) vectors for the \( 2m \) different maturities we obtain the \((2m \times 1)\)

\[ \text{Alternatively, as in Fama (1984) the following regression can be performed:} \]

\[ f_t(t, T) - e_{hf}(T) = \beta_3 + \beta_4 (f_t(t, T) - e_{hf}(t)) + \text{residual}, \]
vector \( f_\theta(\theta e T) \), where \( E [f_\theta(\theta e T)] = 0 \). As described above the GMM procedure consists of replacing \( E [f_\theta(\theta e T)] = 0 \) by its sample counterpart, \( g_\theta (\theta e T) \), using the \( n \) observations,

\[
g_\theta (\theta e T) = \frac{1}{n} \sum_{t=1}^{n} f_\theta (\theta e T, X_t).
\]

From our theoretical model we know that the Fama regression slope parameter, \( \beta_{2T} \), is defined as

\[
\beta_{2T} = \frac{\text{cov} [q_T, r_{pT} + q_T]}{\text{var} [r_{pT} + q_T]} < 0
\]

(6.28)

where \( q_T \) and \( r_{pT} \) are defined as in equations (5.34) and (5.35). Since for short maturities, that is for \( \tau \to 0 \), we have that \( K_{T \tau} \to R_{4,t} \), and therefore we obtain for the spot exchange rate

\[
\text{cov} [q_T, R_h - R_f] = \xi_{ct} [\lambda^*_h - \lambda^*_f] \text{var} [Y] + \gamma^*_h \alpha_{z_h} \text{var} [X_h] + \gamma^*_f \left( \alpha_{z_f} - \sigma^2_{n_{z,f}} \right) \text{var} [X_f]
\]

(6.29)

and

\[
\text{var} (R_h - R_f) = \gamma^2_f \text{var} (X_f) + \gamma^2_h \text{var} (X_h) + [\lambda^*_h - \lambda^*_f]^2 \text{var} (Y).
\]

(6.30)

For the longer maturities we obtain the following expression

\[
\text{cov} [q(t,T), r_{pT} (t,T) + q (t,T)] = B^*_y (T,t) \left[ B_{y_h} (t,T) - B_{y_f} (t,T) \right] \text{var} [Y] + B^*_z (T,t) B_{z_h} (t,T) \text{var} [X_h] + B^*_z (T,t) B_{z_f} (t,T) \text{var} [X_f]
\]

(6.31)

and
\[ \text{var}(rp(t,T) + q(t,T)) = B_{Z_f}^2(t,T) \text{var}(X_f) + B_{Z_h}^2(t,T) \text{var}(X_h) + [B_{\eta_h}(t,T) - B_{\eta_f}(t,T)]^2 \text{var}(Y). \] (6.32)

The $B_i$ and $B_i^*$ terms are defined as in Chapter 5. From the unconditional second moment of the gamma distribution we obtain \( \text{var}(X_h) = \theta_{zh}\sigma_{zh}^2/2\kappa_{zh} \), \( \text{var}(X_f) = \theta_{zf}\sigma_{zf}^2/2\kappa_{zf} \), and \( \text{var}(Y) = \theta_{y}\sigma_{y}^2/2\kappa_{y} \). Given the complexity of the slope coefficients ($\beta_{2T}$) and the great number of unknowns in the parameter space $\Theta$, we use the moments of the underlying equilibrium variables to identify some of these unknowns. This also ensures that the calibrated model parameters not only accounts for the currency puzzle but is also consistent with the time-series properties of the underlying variables.

To determine the moments of the equilibrium nominal interest rates, the rate of depreciation of the currency, and the volatility of the changes in the nominal interest rates and the currency price, we use a discrete-time econometric specification of the continuous-time specification [see equations (5.4), (5.6), (5.7), (5.20), and (5.19)] in Chapter 3. The discrete-time specification for the nominal interest rates, the volatility of the changes in the interest rates and exchange rate, and the changes in the spot currency price can be written, respectively, as follows:

\[
R_{h,t} = \mu_{R_h} + \gamma_h X_{h,t} + \lambda_h Y_t, \quad (6.33)
\]
\[
R_{f,t} = \mu_{R_f} + \gamma_f X_{f,t} + \lambda_f Y_t, \quad (6.34)
\]
\[
V_{\Delta R_h}(X_t, Y_t) = \lambda_h^2 \sigma_{y}^2 Y_t + \gamma_h^2 \sigma_{zh}^2 X_{ht}, \quad (6.35)
\]
\[
V_{\Delta R_f}(X_t, Y_t) = \lambda_f^2 \sigma_{y}^2 Y_t + \gamma_f^2 \sigma_{zf}^2 X_{ft}, \quad (6.36)
\]
\[
V_{\Delta e}(X_t, Y_t) = \sigma_{zh}^2 X_h(t) + \sigma_{zf}^2 X_f(t) + \sigma_{\eta_y}^2 Y(t), \quad (6.37)
\]
\[
e_{t+1} - e_t = \alpha_{zh} X_{ht} - \left(\alpha_{zf} - \sigma_{zf}^2\right) X_{ft} + \xi_{t+1} Y_t + e_{e,t+1}, \quad (6.38)
\]

where

\[
e_{e,t+1} = \sigma_{zh} X_{ht}^{1/2} e_{e,t+1} - \sigma_{zf} X_{ft}^{1/2} e_{e,t+1} + \sigma_{\eta_y} Y_t^{1/2} e_{e,t+1}.
\]
\[ \epsilon_{e_{i,t+1}} \sim (0, 1) \quad i = 1, 2, 3 \]
\[ \text{Cov} (\epsilon_{e_{i,t+1}}, \epsilon_{e_j,t+1}) \neq 0 \quad \text{for } i \neq j. \]

From the system of equations [(6.33) - (6.38)] we use the following moments: the unconditional mean of all variables; the unconditional variances of both nominal interest rates; and the unconditional co-variances between the first five variables. We exclude \( \text{Cov}(V_{\Delta R_j}, V_{\Delta e}) \) from the system, as it leads to a singular matrix of moment equations. These moments are estimated as a set of orthogonality conditions in a GMM framework as described above. We use the one-month Eurodollar interest rates, the one-month Euro-pound interest rate, the spot dollar price of the UK-pound, the volatility series of the interest rates generated in Section 6.2.

To obtain estimates of the unobservable volatility of the currency price dynamics, we use the same GARCH-framework as applied in the case of the interest rates volatilities in section 6.2. To estimate the volatility of currency price changes we model, as in Longstaff and Schwartz (1993), discrete changes in the spot exchange rate by the following econometric specification

\[ \epsilon_{i+1} - \epsilon_i = a_0 + a_1 \epsilon_i + a_2 V_{t,i} + \epsilon_{i,t+1}, \]

where \( \epsilon_{i,t+1} = z_{\epsilon,t+1} \sqrt{V_{t,i+1}} \) is the prediction error of the rate of depreciation and \( \epsilon_{i,t} \sim N(0, V_{t,i}) \). As before the error term, \( z_{\epsilon,t+1} \), is normalized to have a variance of one. Based on the 'joint sign and size test', which confirms the presence of asymmetric GARCH\(^4\), we use the EGARCH specification to estimate the volatility of changes in the spot currency price. The extended Nelson's (1991) asymmetric representation of the

\[ 4 \epsilon_t^2 = 0.00 - 0.003 \times S_{t-1}^- + 0.002 \times S_{t-1}^+ \epsilon_{t-1} + 0.002 \times S_{t-1}^+ \epsilon_{t-1} + 0.002 \times S_{t-1}^+ \epsilon_{t-1} \text{ and the Wald test for the joint hypothesis of zero on the size and sign parameters reports a } P\text{-value} = 0.000. \text{ (The t-statistics are in brackets.)} \]
conditional variance of the rate of depreciation is as follows,

\[
\ln (V_{t,t}) = b_0 + b_1 e_{t-1} + b_2 \ln (V_{t,t-1}) + b_3 (|z_{t-1}|) + b_4 z_{t-1}.
\] (6.40)

The parameter estimates for the EGARCH specification (6.40) confirm the empirical results for this type of specification [see Table 6]. The estimate for \( b_4 \) is negative, which reflects the presence of leverage effect in the non-linear specification. In addition it can be noted from the results that \(-1 < b_4 < 0\), which imply, as noted before, that positive shocks (appreciations) have a smaller effect on the exchange rate volatility than negative shocks (depreciations).

Given these estimates of the moments and the regression slope coefficients we can solve for the model parameters represented in \( \Theta \) as follows. By using the unconditional moments of the state variables, \( E(X_t) = \theta_x, E(Y) = \theta_y, \) \( \text{Var}(X_t) = \theta_x \sigma^2_{z,t}/2\kappa_x, \) and \( \text{Var}(Y) = \theta_y \sigma^2_{v}/2\kappa_y \) we obtain the following system of moment equations:

\[
\begin{align*}
E(R_{h,t}) &= \mu_h + \gamma^*_h \theta_{z_h} + \lambda^*_h \theta_y \\
E(R_{f,t}) &= \mu_f + \gamma^*_f \theta_{z_f} + \lambda^*_f \theta_y \\
E(V_{\Delta h,t}) &= \lambda^*_h \sigma^2_{v} \theta_{y} + \gamma^*_h \theta_{z_h} \\
E(V_{\Delta f,t}) &= \lambda^*_f \sigma^2_{v} \theta_{y} + \gamma^*_f \theta_{z_f} \\
E[V_{\Delta x,t}] &= \sigma^2_{z_h} \theta_{z_h} + \sigma^2_{z_f} \theta_{z_f} + \sigma^2_{\theta_y} \theta_y \\
E(e_{t+1} - e_t) &= \alpha_{z_h} \theta_{z_h} - \xi_{z} \theta_{z_f} + \xi_{e} \theta_{y} \\
\text{Var}(R_{h,t}) &= \gamma^*_h \theta_{z_h} \sigma^2_{z} + \lambda^*_h \theta_{y} \sigma^2_{v} / 2\kappa_x \\
\text{Var}(R_{f,t}) &= \gamma^*_f \theta_{z_f} \sigma^2_{z} + \lambda^*_f \theta_{y} \sigma^2_{v} / 2\kappa_y \\
\text{Cov}(R_{h,t}, R_{f,t}) &= \lambda^*_h \lambda^*_f \theta_{y} \sigma^2_{v} / 2\kappa_y \\
\text{Cov}(R_{h,t}, V_{\Delta h,t}) &= \lambda^*_h \lambda^*_f \sigma^2_{v} \theta_{y} / 2\kappa_y 
\end{align*}
\] (6.41)
Solving this system of equations together with equations (6.28) - (6.32) give us the calibrated values for the 19 unknowns in the parameter space $\Theta$. This solution is consistent with both the currency puzzle and the time-series properties of the data. In the next section we examine the calibrated parameter values.

### 6.4 The empirical results

#### 6.4.1 The dynamics of the nominal interest rates, the rate of depreciation and the term structure

In this section we examine whether the structure of our general equilibrium model is supported by the data. To achieve this validation we investigate whether the model for the dynamics of the nominal interest rates and the rate of depreciation of the spot exchange rate obtained in equations (5.5) and (5.19) provides a good description of the properties of the monthly observations on these variables. In addition, we review whether the cross-sectional restriction implied by the affine term structure model is supported by the data.
Table 4 reports the parameter estimates and the GMM minimized criterion values for the discrete-time model in Equations (6.18), (6.19) and (6.20). The estimates of our model support several features of the short-term riskless interest rate dynamics and the rate of depreciation of the rate of exchange. First, our model describes the dynamics of the variables in question to be driven by a factor structure that consists of one common and two country specific state variables. These state variables can be mapped into the short-term interest rate differential between countries and the volatility of the changes of the interest rates. The GMM minimized criterion $\chi^2$ values, which provides a measure of goodness-of-fit of the model, do not reject the model at conventional significance level, i.e. one percent level. These results show that the three state variables, $R_{diff,t}$, $V_{rR,t}$, and $V_{Rf,t}$, not only have explanatory power for the dynamics of the short-term interest rates and the rate of depreciation of the exchange rate, but that they have it in the way implied by the model. This estimation result means that both the conditional mean and the conditional volatility of the changes in these variables are state dependent.

Second, the mean-reverting property of the short-term interest rates model is also supported by the data. This is important because, as argued by Chan et al. (1992), this feature of the short-term interest rate dynamics makes the term structure models so complex. Therefore, it is important to know whether it plays a significant role in describing the short-term rate process. The parameter estimates for $a_1$, $a_2$, and $a_3$ are significantly different from zero. In contrast, studies of Chan et. al. (1992) found weak evidence of mean reversion in the short-term interest rate. Although the empirical study of Koedijk et al. (1997) focuses on the dynamics of short-term interest rate volatility, it can also be noted in their study that the parameter estimates of the mean reversion for several models of interest rate dynamics are not significant.

Furthermore, the parameters estimates, as reported in Table 4, indicate that the expected rates of change of both US and UK short-term riskless interest rates are positively affected by the interest rate differential. The volatility of the US short-term interest rate has a decreasing impact on both interest rate dynamics, while the volatility of the UK rate
has opposite affect. The expected rate of depreciation of the UK-pound is significantly determined by the short-term interest rate differential and the volatility of the US short-term interest rate. The results for our model’s specification [see equation (6.20)] shows that the UIP-condition is indeed not supported by the data. The impact of the interest rate differential is still negative, but the parameter estimate is smaller than the estimates of the UIP-relation (see Table 1). The explanation therefore is that the volatility of the two interest rates is also incorporated in the expected rate of depreciation. Note that the equilibrium relation we have estimated implies that the expected rate of change of the short-term interest rate fluctuates around its long-run mean, the interest rate differential. Since a relation between co-integrated variables can be represented by an error-correction mechanism, the estimated interest rate can be re-written in an error-correction form.

Finally, we also find that the conditional volatility of the interest rate process and the rate of depreciation are highly sensitive to the interest rate differential and the interest rate volatilities. The parameter estimates indicate that an increase in the interest rate differential increases the volatility of the interest rate processes. It is obvious that an increase in the volatility of the interest rates has a positive impact on the volatility in the next period. It is not clear why the volatility of the US-interest rate has a negative impact on the volatility of UK-euro-currency interest rate. The volatility of the rate of depreciation of the UK-pound is positively affected by the volatility of the short-term UK-interest rate. As expected both the interest rate differential and the volatility of the US interest rate has a decreasing impact on the volatility of the depreciation of the UK-currency price.

Next, we review the estimation results of the cross-sectional restriction of the terms structure of interest rate model developed in Equation (6.24). Table 5 reports the GMM estimates of the parameter values, their corresponding t-statistics, and the $R^2$ for each equation. The cross-sectional restrictions of the model are rejected when the GMM minimized criterion value is very large. The value for this $\chi^2_{24}$ test statistic is 0.0816, which is smaller than the critical value at conventional significance levels. Hence, the
cross-sectional restrictions imposed by the three-factor, two-country term structure model cannot be rejected by the data.

Since we have estimated the cross-sectional restrictions of the term structure model in a linearized form we can compare the parameter estimates across maturities. A salient feature of our result is that changes in the interest rate differential have a significant impact on yield changes, and that this affect is decreasing with maturity. For instance, the estimated parameter for changes in the interest rate differential is 0.606 with respect to changes in the short end of the US-yield curve. It drops unevenly to 0.322, 0.385, and 0.331 for the three- and six-month and the one-year yields. As expected, changes in the interest rate differential have opposing effects on US- and UK-yield changes. The impact on the dynamics of US-yields is positive for all maturities, while negative for the UK-yield changes.

Another remarkable result is that the US-yield dynamics is sensitive to volatility changes of the short-term riskless US-interest rate only at the short end of the yield-curve. Volatility changes of the short-term US-interest rate have a smaller and insignificant effect on the longer maturities. In contrast, the volatility changes of the short-term UK-interest rate dynamics have a significant impact on the dynamics of the US-yield curve across all maturities, although the magnitude of the impact also decreases with longer maturities. All these features indicate that short-term maturities are more sensitive to short term dynamics, than the longer-term maturities.

A review of the estimation results for the dynamics of the UK-yields shows that changes in the interest rate differential have a significant impact on yield changes. Note that while the magnitude of this effect decreases with maturity, this decrease is not as pronounced as in the case of the US-yields. In contrast to the US-yields, changes in the UK-yields are significantly more sensitive to volatility changes in the short-term US-interest rate dynamics. Except for the six-month yield, which is not significant, the magnitude of impact of these volatility changes remain more or less constant across the different UK-maturities in the test. The effect of the volatility changes on the UK-yields
dynamics is in conformity with the US-yields. It has a significant effect on the dynamics of the UK-yield curve, except for the one-year yield changes. Again, the magnitude of this impact is decreasing with maturity.

We conclude that the structure of our general equilibrium model is supported by the data, since both the dynamic properties as the cross-sectional restrictions of the equilibrium quantities are validated by the empirical evidence.

6.4.2 Accounting for the currency puzzle

Table 7, Panel I, presents the estimation results for the linear specification of the term structure of foreign exchange returns. We regressed the expected foreign exchange return on the yield differential, observed at time $t$ with maturities similar to that of the forward foreign exchange contract. We use three different maturities in our empirical test, namely one month, six months, and one year. The GMM-minimized criterion value is zero as the system is exactly identified. The system of equations provides us with three Fama slope parameter estimates, namely $-2.631$, $-2.009$, and $-1.394$ for, respectively, the one month, six months, and one year maturities. These point estimates are significantly different from one, at conventional level of significance. We can observe that all the forward premium terms enter with a negative coefficient, which confirm the existence the forward premium puzzle. This result implies that, when the US-dollar interest rates are high relative to the UK-pound rates at any point in the term structure, there is a tendency for the US-dollar to appreciate vis-a-vis the UK-pound. In addition it is noticeable that the magnitude of the estimated slope coefficient in absolute terms is decreasing with maturity of the yields, i.e. the differential at the longer maturities has a smaller effect on the rate of depreciation of the currency. As is expected, the dynamics of the exchange rate are largely determined by developments at the short end of the term structure of interest rates. Clarida and Taylor (1997), in a different context, report similar result in the case of the dollar-mark and the dollar-yen rates, but with the opposite signs.

We determine the parameter values of our model such that they exactly match these
point estimates of the slope parameters of the term structure of foreign exchange return regression and the moments of the underlying equilibrium variables reported in Table 7, Panel II. The calibration result is presented in Table 7, Panel III. Before examining the economic implication of the computed model parameter values, we consider the plausibility of these values and the implication for the Fama condition. Fama (1984) formulated the following conditions in order to account for the currency puzzle: (a) negative covariance between \( r_p(t,T) \) and \( q(t,T) \) [see equations (5.34) and (5.35) for definitions] and (b) greater variance of \( r_p(t,T) \) than \( q(t,T) \). Fama's first condition is satisfied by construction, since we calibrate the parameter values to match the (negative) point estimates of the slope parameters. Recall that the latter is defined as 
\[
\beta_{2T} = \frac{\text{cov}[q_T, r_p_T] + \text{var}[q_T]}{\text{var}[r_p_T]}
\]  
[see equation (6.28)]. As can be observed from Table 8 below the calibrated parameter values also satisfy the second Fama-condition, as the variability of the forward risk premium is larger than that of the expected exchange rate return for all maturities. Noteworthy is the sharp increase in the variability at longer maturities, which seems intuitive given the increased uncertainty at the long end of the term structure.

<table>
<thead>
<tr>
<th>Table 8 Fama's second condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturities</td>
</tr>
<tr>
<td>1-month</td>
</tr>
<tr>
<td>\text{Var}(r_p(t,T))</td>
</tr>
<tr>
<td>\text{Var}(q(t,T))</td>
</tr>
</tbody>
</table>

Since we calibrate the parameter values to match the moments of the time series in question (the spot US and UK interest rates, the rate of depreciation of the dollar price of the UK-pound, and their volatility series), it partially guarantees the plausibility of these values. The parameter that represents the autonomous expected level of interest rates, \( \mu_{R_t} \), is in accordance with the observed first moment conditions for the level of
interest rates in both countries, that is the long term unconditional mean of the UK rate is larger than that of the US rate. This is also reflected by the parameter values for $\mu_{R_i}$, for $i \in \{\text{US, UK}\}$. The fact that $\mu_{R_{i,K}}$ is more than twice as large as $\mu_{R_{i,U}}$ is possibly due to the larger volatility of the US interest rate.\footnote{Note that this term is defined as $\mu_{R_i} = \rho + \mu_m - \sigma_m^2$.} Note also that the long term mean of the local UK-state variable ($\theta_{z_{UK}}$) is also larger than its US counterpart.

A more fundamental condition in this context is that the parameter values guarantee the non-existence of negative nominal interest rates in both countries, since negative interest rate is the trade-off of accounting for the currency puzzle in these type of term structure models [see Backus et al. (2001)]. In term structure of interest rate models this condition is usually referred to as the Feller (1951) condition. The Feller (1951) condition implies that for the stochastic process in equation (5.5) to preclude a positive probability of negative nominal interest rates in both countries, the mean must be larger than the variance, that is $2\kappa_i \theta_{z_i} + 2\kappa_i \theta_{y_i} > \gamma_i^* \sigma_{z_i}^2 + \lambda_i^* \sigma_{y_i}^2$, for $i \in \{\text{US, UK}\}$. By using the parameter values in Table 7, Panel III we can observe that the stochastic processes governing the equilibrium nominal interest rates in equation (5.5) satisfies this condition, despite the fact that the diffusion parameter that captures real local shocks in the US economy ($\sigma_{z_{US}}^2$) is relatively large. In addition, we can observe that the value of the parameter that captures the speed of adjust to real local shocks in the UK economy ($\kappa_{z_{UK}}$) is also relatively small. The relatively small parameter value of the monetary response parameters in both countries ($\gamma_{US}^*$ and $\gamma_{UK}^*$) compensate for these extreme values.

The economic interpretation of the estimated parameter values is as follows. From the relation between the $\text{Cov}(R_{US,t}, V_{\Delta R_{UK,t}})$ and $\text{Cov}(R_{UK,t}, V_{\Delta R_{US,t}})$ in equations (6.41j) and (6.41k), respectively, and their sample moments in Table 7, Panel II we can observe that $\lambda_{U_K}^* = 0.66 \lambda_{U_S}^*$. This relation basically implies that the monetary response is larger in the US and that a common, say international, shock has, therefore, a larger affect on the US-interest rates than on the UK-interest rates. Thus, for instance, a positive
international shock increases the interest rate differential between these countries, while a negative shock reduces this differential. This holds for all maturities. At the other hand, the negative value for $\alpha_y = \alpha_{yh} - \alpha_{yf} - (\sigma_{\eta_yh} - \sigma_{\eta_yf}) \sigma_{\eta_yf}$ entails that the same shock has probably a larger effect on the real rate of production in the UK than in the US ($\alpha_{yh} < \alpha_{yf}$). For example an international economic shock causes the real rate of return on productive assets in the UK to decline relative to the US return. To ensure that investors remain indifferent between investment in these two countries the US-dollar depreciates in value vis-a-vis the Uk-pound, such that the rate of return expressed in their respective currencies remains equal. The negative international shock, due to the relatively large response in the US, decreases the US rate versus the UK rate. The reduced interest rate differential between the US and the UK combined with the depreciation of the dollar price of the UK-pound explains the negative co-variability that underlies the Fama puzzle. The same analysis applies for longer maturities as the $B^*_y(T, t)$ terms are negative, namely $B^*_y(6, t) = -0.5$ and $B^*_y(12, t) = -0.9$, while $B_{YuU}(t, T) > B_{YuK}(t, T)$ for both maturities considered in the sample. The negative $B^*_y(T, t)$ terms implies that, for instance, international developments have a stronger impact on the long term rate of productive returns in the UK than in the US. In contrast, due to the larger monetary response in the US, it tends to reduce the bond price in the US relative the UK. The data fully supports the arguments put forward that the asymmetric affect of shocks on the real rate of return in production and on the riskless rate of return on nominal assets explains the negative co-variability observed in the data, i.e. the Fama puzzle.