DISCUSSION PAPER NO. 404

The Term Structure of the Forward Premium

November 15, 1979

by

Craig S. Hakkio

Northwestern University
and
National Bureau of Economic Research

I. Introduction

There have been numerous studies of the efficiency of the foreign exchange (see Levich 1979 for a survey of many of these studies). However, most of these studies focus on a single maturity - usually a one month forward exchange rate. We observe that forward contracts of many maturities are simultaneously traded in the foreign exchange market; yet there are surprisingly few studies that examine the implications of several forward contract maturities (see Porter 1971, Giddy 1977 and Brillembourg 1978). The hypothesis of market efficiency has well known implications for the relation between a forward exchange rate of a given maturity and the subsequently observed spot rate. In addition, the hypothesis of efficiency has implications for the joint behavior of forward exchange rates of various maturities. This paper will theoretically and empirically examine these additional implications.

Section II will propose an equilibrium theory of the term structure of the forward premium. By combining the (certainty equivalence) theory of the term structure of (domestic and foreign) interest rates with the hypothesis of interest rate parity, a simple expression relating the six month forward premium to the expected future one month forward premium can be derived. It will be shown that the six month forward premium can be written as a geometric average of expected future one month forward premiums. In Section III it is shown that a convenient and efficient method to extract the expected one month forward premium can be obtained by assuming that a general (bivariate) stochastic process generates the one and six month forward premiums. The theory developed in Section II will then impose highly nonlinear cross equation restrictions on the parameters of the stochastic process.

The restrictions imposed on the parameters of the model by the economic theory are highly nonlinear. Sections IV and V discuss two methods of testing the validity of the restrictions. Section IV provides a statistical test of the hypothesis that requires only the unrestricted estimates. The rejection region, under the null hypothesis, is derived. The statistical test proposed in Section V requires the restricted parameter estimates. Maximum likelihood methods for estimating the constrained models are discussed and implemented.

II. The Economics of the Term Structure of the Forward Premium

To develop a theory of the term structure of the forward premium, we begin by assuming that interest rate parity holds. (See Porter 1971 for a similar development). There is much empirical evidence in support of this condition (Frenkel and Levich 1975, 1977). Interest rate parity states that the expected rate of depreciation on foreign exchange equals the interest rate differential. We can write this as:

$$\frac{1 + I_{n,t}}{1 + I^{*}_{n,t}} = \frac{E_t S_{t+n}}{S_t} \tag{1}$$

$$1 + i_{n,t} = \frac{(1 + I_{n,t})^{n}}{(1 + I_{n-1,t})^{n-1}} \tag{2a}$$

$$(1 + I_{n,t})^{n} = (1 + i_{1,t})(1 + i_{2,t}) \cdots (1 + i_{n,t}) \tag{2b}$$

where $I_{n,t}$ = n period rate of interest at period t

$i_{n,t}$ = implicit one period forward interest rate for period t + n.

Dividing by the foreign country version of (2a), (2a*), yields:

$$\frac{1 + i_n}{1 + i_n^*} = \left[\frac{1 + I_n}{1 + I_n^*}\right]^n \Big/ \left[\frac{1 + I_{n-1}}{1 + I_{n-1}^*}\right]^{n-1} \tag{3}$$

Substituting interest rate parity (1) into (3) and cancelling yields:

$$\frac{E_t S_{t+n}}{E_t S_{t+n-1}} = \frac{1 + i_n}{1 + i_n^*} = 1 + E_t \bar{r}_{1,t+n-1} \tag{4}$$

$E_t \bar{r}_{1,t+n-1}$ is the implied expected change in the spot exchange rate in period t+n. Dividing (2b) by (2b*), equating to (3) and substituting from (4) yields:

$$\frac{E_t S_{t+n}}{S_t} = \frac{(1 + i_1)(1 + i_2) \ldots (1 + i_n)}{(1 + i_1^*)(1 + i_2^*) \ldots (1 + i_n^*)} \tag{5}$$

$$= (1 + \bar{r}_{1,t})(1 + E_t \bar{r}_{I,t+1}) \ldots (1 + E_t \bar{r}_{1,t+n-1}).$$

Define

$$r_{k,t} = \left[\frac{F_{k,t}}{S_t}\right]^{1/k} - 1. \tag{6}$$

where $F_{k,t}$ is the k-month forward exchange rate prevailing at time t, and $r_{k,t}$ is the k-month forward premium. If we assume that the forward rate is an unbiased predictor of the future spot rate, $F_{k,t} = E_t S_{t+k}$, then we can conclude that the k-month forward premium, $r_{k,t}$, equals the expected k-month rate of depreciation, $E_t \bar{r}_{t+k}$. Therefore, substituting the definition of $r_{k,t}$ from equation (6), for k = 1 and n, into equation (5) we obtain

$$(1 + r_{n,t})^{1/n} = (1 + r_{1,t}) \cdots (1 + E_t\, r_{1,t+n-1})$$

or, as an approximation (using $\ell n\,(1 + x) = x$, for small x):

$$r_{n,t} = \frac{1}{n}\,[r_{1,t} + E_t r_{1,t+1} + \cdots + E_t r_{1,t+n-1}] \tag{7}$$

E_t is the mathematical expectations operator conditional on the set of information available to economic agents at time t, Ω_t. We will assume that $\Omega_t \supseteq \Omega_{t-1} \supseteq \cdots$, and that Ω_t contains at least all current and lagged values of $(r_{1,t}, r_{n,t})$. The following derivation follows Sargent's (1979) analysis of the term structure of interest rates.

To add empirical content to equation (7), we must specify how expectations are formed and what variables belong in Ω_t. First, notice that the conditional expectations operator in (7), E_t, has Ω_t as the conditioning set, where Ω_t includes all relevant information for calculating expectations of future one month rates of depreciation. For convenience in deriving testable implications based on (7), let us write $E_t x_s$ as $E(x_s|\Omega_t)$, and so rewrite (7) as

$$r_{n,t} = \frac{1}{n}\,[r_{1,t} + E(r_{1,t+1}|\Omega_t) + \cdots + E(r_{1,t+n-1}|\Omega_t)] \tag{7'}$$

Let θ_t be any subset of Ω_t, such that θ_t includes at least current and lagged values of $r_{1,t}$ and $r_{n,t}$. Now, take expectations of both sides of (7'), conditional on the <u>smaller</u> information set θ_t, to obtain

$$r_{n,t} = \frac{1}{n}\,[r_{1,t} + E(r_{1,t+1}|\theta_t) + \cdots + E(r_{1,t+n-1}|\theta_t)] \tag{7''}$$

where we used the law of iterated projections that states that $E(y|z) = E\{E(y|x,z)|z\}$, where x, y, z are normal random variables. Notice that (7'') and (7') are of the same form. In particular, if we leave important variables out of θ_t (that were in Ω_t) we will not invalidate the tests reported. Now, take first differences of (7''):

$$r_{n,t} - r_{n,t-1} = \frac{1}{n}\{(r_{1,t} - r_{1,t-1}) + [E(r_{1,t+1}|\theta_t) - E(r_{1,t}|\theta_{t-1})] + \ldots + [E(r_{1,t+n-1}|\theta_t) - E(r_{1,t+n-2}|\theta_{t-1})]\}. \tag{8}$$

Write $\Delta r_{n,t} = r_{n,t} - r_{n,t-1}$ and $\Delta r_{1,t} = r_{1,t} - r_{1,t-1}$. Take expectations of both sides of (8), conditional on θ_{t-1}, using the law of iterated projections, to get

$$E(\Delta r_{n,t}|\theta_{t-1}) = \frac{1}{n}\{E(\Delta r_{1,t}|\theta_{t-1}) + E(\Delta r_{1,t+1}|\theta_{t-1}) + \ldots + E(\Delta r_{1,t+n-1}|\theta_{t-1})\}. \tag{8'}$$

We must now specify <u>exactly</u> which variables to include in θ_t. We shall restrict θ_t to include only current and lagged values of $\Delta r_{1,t}$, $\Delta r_{n,t}$, that is, $\theta_t = \{\Delta r_{1,t}, \Delta r_{1,t-1}, \ldots, \Delta r_{n,t}, \Delta r_{n,t-1}, \ldots\}$. Given this information set, we can easily calculate the conditional expectations in (8'). We shall report two methods of calculating these expectations, and the restrictions implied by (8').

III. The Empirical Implications of the Term Structure of the Forward Premium

Assuming that $(\Delta r_{1,t}, \Delta r_{n,t})$ is a linearly indeterministic covariance stationary stochastic process, we can use the Wold Decomposition Theorem to write (letting $R_{1,t} = \Delta r_{1,t}$ and $R_{n,t} = \Delta r_{n,t}$):

$$
\begin{aligned}
R_{1,t} &= \alpha(L)w_t + \beta(L)v_t \\
R_{n,t} &= \gamma(L)w_t + \delta(L)v_t
\end{aligned}
\tag{9}
$$

where $\alpha(L)$, $\beta(L)$, $\gamma(L)$ and $\delta(L)$ are one sided polynomials in the lag operator

$w_t = R_{1,t} - E(R_{1,t}|\theta_{t-1})$

$v_t = R_{n,t} - E(R_{n,t}|\theta_{t-1})$

$$Ew_t w_{t-k} = \begin{cases} \sigma_w^2 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

$$Ev_t v_{t-k} = \begin{cases} \sigma_v^2 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

$$Ew_t v_{t-k} = \begin{cases} \sigma_{wv} & k = 0 \\ 0 & k \neq 0 \end{cases}$$

$\alpha(0) = \delta(0) = 1$ and $\beta(0) = \gamma(0) = 0$.

The Weiner-Kolmogorov prediction formulas allow us to write the conditional expectations in (8') in a simple fashion:

$$E_{t-1}R_{1,t+k} = \left[\frac{\alpha(L)}{L^{k+1}}\right]_+ w_{t-1} + \left[\frac{\beta(L)}{L^{k+1}}\right]_+ v_{t-1} \tag{10}$$

where $[\ ]_+$ means "ignore negative powers of L." Substituting expression (10) into (8') and rearranging, yields:

$$E_{t-1}R_{n,t} = \frac{1}{n}\left[\frac{\alpha(L)}{L} + \frac{\alpha(L)}{L^2} + \dots + \frac{\alpha(L)}{L^n}\right]_+ w_{t-1}$$

$$+ \frac{1}{n}\left[\frac{\beta(L)}{L} + \frac{\beta(L)}{L^2} + \ldots + \frac{\beta(L)}{L^n}\right]_+ v_{t-1} \tag{11}$$

$$= \frac{1}{n}\left[\frac{\alpha(L)}{L}\left(\frac{1-L^n}{1-L^{-1}}\right)\right]_+ w_{t-1} + \frac{1}{n}\left[\frac{\beta(L)}{L}\left(\frac{1-L^{-n}}{1-L^{-1}}\right)\right]_+ v_{t-1}.$$

But, we can also use the Weinter-Kolmogorov prediction formula to write the left hand side of (11) as

$$E_{t-1}R_{n,t} = \left[\frac{\gamma(L)}{L}\right]_+ w_{t-1} + \left[\frac{\delta(L)}{L}\right]_+ v_{t-1}. \tag{12}$$

Equating terms in (11) and (12) yield a set of cross equation restrictions on the parameters of the bivariate moving average representation of $(R_{1,t}, R_{n,t})$ in (9) implied by the theory of the term structure of the forward premium:

$$\begin{aligned} \left[\frac{\gamma(L)}{L}\right]_+ &= \frac{1}{n}\left[\frac{\alpha(L)}{L}\left(\frac{1-L^{-n}}{1-L^{-1}}\right)\right]_+ \\ \left[\frac{\delta(L)}{L}\right]_+ &= \frac{1}{n}\left[\frac{\beta(L)}{L}\left(\frac{1-L^{-n}}{1-L^{-1}}\right)\right]_+ . \end{aligned} \tag{13}$$

It is possible to estimate equation (9) subject to (13) and so test the validity of the restrictions embodied in (13). However, it is very difficult to estimate constrained bivariate moving averages and so we use an alternative representation of $(R_{1,t}, R_{n,t})$.

Equation (8') is a restriction across the systematic part of $(R_{1,t}, R_{n,t})$, imposed by rational expectations. By our assumptions of stationarity we can write $(R_{1,t}, R_{n,t})$ as a vector autoregression (where the α's, β's, γ's and δ's in (14) are different than those in (9), the w_t and v_t are the same):

$$R_{1,t} = \sum_{i=1}^{M} \alpha_i R_{1,t-i} + \sum_{i=1}^{M} \beta_i R_{n,t-i} + w_t \tag{14a}$$

$$R_{n,t} = \sum_{i=1}^{M} \gamma_i R_{1,t-i} + \sum_{i=1}^{M} \delta_i R_{n,t-i} + v_t \tag{14b}$$

where $Ew_t R_{1,t-i} = Ev_t R_{1,t-i} = Ew_t R_{n,t-i} =$

$$Ev_t R_{n,t-i} = 0, \quad \text{for } i = 1, 2, \ldots, M$$

$$Ew_t v_{t-i} = \begin{cases} 0 & i \neq 0 \\ \sigma_{wv} & i = 0 \end{cases}$$

$\{w_t, v_t\}$ is the innovation in the $(R_{1,t}, R_{n,t})$ process; the errors are contemporaneously correlated, but uncorrelated at all lags. Equation (14) can be rewritten as:[1]

$$x_t = Ax_{t-1} + a_t \tag{15}$$

where

$$x_t = \begin{vmatrix} R_{1,t} \\ R_{1,t-1} \\ \cdot \\ \cdot \\ \cdot \\ R_{1,t-M+1} \\ R_{n,t} \\ R_{n,t-1} \\ \cdot \\ \cdot \\ \cdot \\ R_{n,t-M+1} \end{vmatrix} \qquad a_t = \begin{vmatrix} w_t \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ v_t \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{vmatrix}$$

---

[1] Equation (15) amounts to rewriting an M^{th} order difference equation as a vector first order system.

$$A = \left|\begin{array}{cccccccccc}
\alpha_1 & \alpha_2 & \cdots & \alpha_{M-1} & \alpha_M & \beta_1 & \beta_2 & \cdots & \beta_{M-1} & \beta_M \\
1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\gamma_1 & \gamma_2 & \cdots & \gamma_{M-1} & \gamma_M & \delta_1 & \delta_2 & \cdots & \delta_{M-1} & \delta_M \\
0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 0
\end{array}\right|
\begin{array}{l}
\\
\\
\\
\\
\longleftarrow \text{ row } M+1 \\
\\
\\
\longleftarrow \text{ row } 2M.
\end{array}$$

Repeated substitution from (15) yields:

$$\begin{aligned}
x_{t+1} &= Ax_t + a_{t+1} = A^2 x_{t-1} + Aa_t + a_{t+1} \\
x_{t+j} &= A^{j+1} x_{t-1} + A^j a_t + \ldots + a_{t+j}.
\end{aligned} \tag{16}$$

Since $E_{t-1} a_{t+k} = 0$ $(k = 0, 1, \ldots)$ we can write (16) as

$$E_{t-1} x_{t+j} = A^{j+1} x_{t-1} \qquad j = 1, 2, \ldots \tag{17}$$

Letting $c = (1\ 0\ .\ .\ .\ 0)$ and $d = (0\ \ldots\ 0\ 1\ 0\ \ldots\ 0)$, we can write

↑ column $M + 1$

$$\begin{aligned}
R_{1,t} &= c\ x_t \\
R_{n,t} &= d\ x_t.
\end{aligned} \tag{18}$$

Multiply (15) by c and d to get

$$c\ x_t = c\ Ax_{t-1} + c\ a_t$$

$$d\ x_t = d\ Ax_{t-1} + d\ a_t.$$

Equating to (18) we get

$$R_{1,t} = c\ Ax_{t-1} + a_{1,t} \qquad (19)$$
$$R_{n,t} = d\ Ax_{t-1} + a_{n,t}.$$

Multiply restriction (17) by c to get

$$E_{t-1}c\ x_{t+j} = c\ A^{j+1}x_{t-1} \quad j = 1,\ 2,\ \ldots \qquad (20)$$

Updating (18) by +j and substituting into (20) yields

$$E_{t-1}R_{1,t+j} = c\ A^{j+1}x_{t-1} \quad j = 1,\ 2,\ \ldots \qquad (21)$$

Substitution of (21) into (8) yields

$$E_{t-1}R_{n,t} = (1/n)(cAx_{t-1} + cA^2x_{t-1} + \ldots + cA^nx_{t-1}) \qquad (22)$$
$$= (1/n)\ c(A + A^2 + \ldots + A^n)x_{t-1}.$$

Taking expectations conditional on I_{t-1} in (19) yields

$$E_{t-1}R_{n,t} = d\ Ax_{t-1}. \qquad (23)$$

Equating equations (22) and (23) yields the following restriction imposed by rational expectations:

$$d\ A = (1/n)\ c(A + A^2 + \ldots + A^n). \qquad (24)$$

The intuition behind these restrictions arise from the following observations. We assumed that the $(R_{1,t},\ R_{n,t})$ process was generated by a vector autoregression. That is, we regress both $R_{1,t}$ and $R_{n,t}$ against lagged values of $(R_{1,t},\ R_{n,t})$. Wold has shown the conditions under which this is valid (see Whittle 1963). If the economic agents realize this they will use the parameters of the autoregression to generate their forecasts. For the data to be consistent with the model the parameter values must be restricted. These restrictions are summarized in equation (8) and equivalently in equation (24).

IV. Econometric Tests and Results, I

The restrictions implied by equation (24) are highly nonlinear. There are two basic methods to test the validity of the restrictions implied by the theory. The first method, discussed in detail in this section, was originally proposed by Wald. This method requires obtaining the unrestricted maximum likelihood estimates $\hat{\psi}^u$ of the parameter vector $\psi = (\alpha, \beta, \gamma, \delta)$. Let us write the restrictions implied by (24) in the form $h(\psi) = 0$. Wald's method then tests $h(\hat{\psi}^u) = 0$. The second method, discussed in detail in the next section is based on the likelihood ratio test. This method requires obtaining in addition to the unrestricted estimate, $\hat{\psi}^u$, the restricted estimate $\hat{\psi}^r$. One then compares the likelihood of $\hat{\psi}^u$ to $\hat{\psi}^r$. A difficulty with this method is obtaining the restricted maximum likelihood estimates. The next section will present two methods of obtaining $\hat{\psi}^r$.

Under the assumption that $\{w_t, v_t\}$ is bivariate normal, the likelihood function for a sample of $\{w_t, v_t\}$, $t = 1, 2, \ldots T$ is given by

$$L(\alpha, \beta, \gamma, \delta) = (2\pi)^{-T} |V|^{-T/2} \exp\{-\frac{1}{2} \sum_{t=1}^{T} e_t' V^{-1} e_t\} \tag{25}$$

where $$e_t = \begin{bmatrix} w_t \\ v_t \end{bmatrix} \quad V = E\, e_t' e_t .$$

Maxmimizing (25) without any restrictions, that is, with all parameters free, is equivalent to estimating (14) by least squares. Wilson (1973) shows that the parameter estimates with an unknown V may be obtained by

$$\min |\hat{V}| = | (1/T) \sum_{t=1}^{T} \hat{e}_t \hat{e}_t' |.$$

To test restriction (24) we proceed as follows. Let

$$\psi = (\underline{\alpha}', \underline{\beta}', \underline{\gamma}'. \underline{\delta}')$$

$$\hat{\psi}^u = \text{OLS (unrestricted) estimate.}$$

Write restriction (24) as

$$h(\psi) = dA(\psi) - (1/n)c(A(\psi) + A^2(\psi) + \ldots + A^n(\psi)) \tag{26}$$

$$= (0\ 0\ \ldots\ 0) = \underline{0}$$

where we write $A = A(\psi)$ to indicate the dependence of A on ψ. The test amounts to testing whether the vector $h(\hat{\psi})$ is significantly different from the zero vector. The problem is to determine the shape of the rejection region. The problem is intuitively solved as follows (Silvey 1975, pp. 115-116, or Rao 1973, pp. 418-419):

We expect $\underline{\psi}^u$ to be "close to" $\underline{\psi}$, under the null hypothesis, and we know that $\sqrt{n}(\underline{\psi}^u - \underline{\psi})$ is approximately $N(0, B_\psi^{-1})$, where B_ψ^{-1} is the information matrix for the coefficient vector $\underline{\psi}$, in a single observation. Expanding $h(\psi)$ about $\underline{\psi}$ in a Taylor series, to linear terms, we get

$$h(\hat{\underline{\psi}}) \simeq h(\underline{\psi}) + H_\psi'(\hat{\underline{\psi}} - \underline{\psi}) \tag{27}$$

where $$H_{\psi} = \left[\frac{\partial h_j(\underline{\psi})}{\partial \psi_i}\right].$$

Since $h(\underline{\psi}) = 0$, under the null hypothesis, we may rewrite (27) as

$$h(\hat{\underline{\psi}}) \simeq H_{\psi}' \, (\hat{\underline{\psi}} - \underline{\psi}). \tag{28}$$

Therefore, $\sqrt{n}\, h(\hat{\psi})$ is approximately $N(0, H_{\psi}' \, B_{\psi}^{-1} \, H_{\psi})$. Letting x be the vector of observations, the rejection region becomes

$$\{x | n[h(\psi(x))'(H_{\psi}' \, B_{\psi}^{-1} \, H_{\psi})^{-1} \, h(\psi(x))] > k\}. \tag{29}$$

To actually apply this test one needs an estimate of H_{ψ} and B_{ψ}^{-1}. For $(1/n)B_{\psi}^{-1}$ we can use the estimated variance-covariance matrix, obtained from estimating (14). For H_{ψ}, we numerically differentiate the 1x8 restriction vector $h(\underline{\psi})$ (at the OLS estimates) with respect to all sixteen parameters. Calling these estimates $(1/n)B_{\hat{\psi}}^{-1}$ and $H_{\hat{\psi}}$, the value W is given by

$$W = h(\psi)'[H_{\hat{\psi}}(1/n)B_{\hat{\psi}}^{-1} \, H_{\hat{\psi}}]h(\psi). \tag{30}$$

Under the null hypothesis, $h(\underline{\psi}) = 0$, W is approximately distributed chi-square with eight degrees of freedom. Large values of W indicate rejection of the hypothesis.

The OLS (unrestricted) estimates are given in Tables 1 and 2, under the heading "Unrestricted Estimates." Also presented is an estimate of V and $|V|$. At the bottom of Tables 1 and 2 the W-statistic is presented, along with its marginal significance level. The marginal significance level is the probability of observing a number greater than the statistic, given that the null hypothesis is true.

The results presented in Table 1 for Germany indicate a failure to reject the validity of the hypothesis that the pure expectations theory of the term structure of the forward premium is correct. The W-statistic of 10.42 is insignificant, as indicated by a marginal significance level of 23.7 percent. The results in Table 2 for Canada indicate rejection of the null hypothesis. The W-statistic of 21.84 is significant, as indicated by a marginal significance level of 0.5 percent. The assumption that $(R_{1,t}, R_{n,t})$ is stationary is equivalent to the assumption that the characteristic roots of the matrix A are all less than one in modulus (see Sargent [1979a], p. 273). The roots of A, using OLS estimates, were calculated and all were found to be less than one in modulus.

V. Econometric Tests and Results, II

In the last section, we presented tests of the validity of restriction (24) based on the unrestricted estimates. In the case of Canada, we cannot determine the source of rejection. In this section, we shall estimate the model with the restrictions imposed, and then compare the restricted and unrestricted models using a likelihood ratio test.

The restrictions implied by (24) are highly nonlinear. Sargent (1979b) proposes two alternative estimation strategies. The first method requires estimating the first row of A, equation (14a), by least squares. Then, the (M+1)st row of A, equation (14b), is calculated using an iterative procedure. Form a preliminary estimate of A, call it A_o, by setting row M + 1 to a row of zeroes, and all other rows to their known (or consistent) values. Calculate the (M+1)st row of A, at iteration i + 1, as

TABLE 1

GERMANY: ESTIMATES OF BIVARIATE AUTOREGRESSION
UNRESTRICTED AND RESTRICTED

j	1	2	3	4
		Unrestricted Estimates		
α_j	-0.3603	0.0726	-0.0856	-0.2098
β_j	0.5599	-0.5044	0.1862	0.4478
γ_j	-0.0767	0.0642	0.0293	-0.0412
δ_j	0.2227	-0.1829	0.0304	0.1210

$$V = \begin{pmatrix} 7.726*10^{-7} & 2.540*10^{-7} \\ & 1.281*10^{-7} \end{pmatrix} \qquad |V| = 3.442*10^{-14}$$

j	1	2	3	4
		Restricted Estimates		
α_j	-0.3651	-0.1119	-0.1779	-0.1117
β_j	0.3003	-0.0845	0.2441	0.2275
γ_j	-0.0797	-0.0410	-0.0264	-0.0098
δ_j	0.0784	0.0447	0.0435	0.0199

$$V = \begin{pmatrix} 7.959*10^{-7} & 2.666*10^{-7} \\ & 1.352*10^{-7} \end{pmatrix} \qquad |V| = 3.652*10^{-14}$$

Likelihood ratio statistic = 11.134

Marginal significance level = 0.194

W = 10.42

Marginal significance level = 0.237

TABLE 2

CANADA: ESTIMATES OF BIVARIATE AUTOREGRESSION
UNRESTRICTED AND RESTRICTED

j	1	2	3	4
		Unrestricted Estimates		
α_j	-0.2830	-0.0174	-0.0168	-0.1421
β_j	0.3743	0.1092	0.1202	0.0028
γ_j	0.1745	0.1186	0.0643	0.0171
δ_j	-0.1140	-0.0249	-0.0161	0.0001

$$\hat{V} = \begin{bmatrix} 9.423*10^{-8} & 4.572*10^{-8} \\ & 4.084*10^{-8} \end{bmatrix} \qquad |\hat{V}| = 1.757*10^{-15}$$

j	1	2	3	4
		Restricted Estimates		
α_j	-0.5183	-0.0885	-0.0077	-0.1151
β_j	0.5196	0.0389	0.0769	-0.0012
γ_j	-0.0673	-0.0227	-0.0135	-0.0098
δ_j	0.0580	0.0118	0.0064	-0.0001

$$\hat{V} = \begin{bmatrix} 9.829*10^{-8} & 4.958*10^{-8} \\ & 4.498*10^{-8} \end{bmatrix} \qquad |\hat{V}| = 1.963*10^{-15}$$

Likelihood ratio statistic = 20.843

Marginal significance level = 0.008

W = 21.84

Marginal significance level = 0.005

$$(\text{row } M{+}1)_{i+1} = dA_{i+1} = c\,\frac{1}{n}\,[A_i + A_i^2 + \ldots + A_i^{M-1}] \tag{31}$$

where A_i is the estimate of A on the i^{th} iteration. At each step in forming A_{i+1}, all rows (except the (M+1)st) are kept equal to the corresponding row of A_o. If this procedure converges, it will find an A that satisfies (24). This procedure will converge if the characteristic roots of A are less than one in modulus. Since the elements of row 1 are consistently estimated by least squares, the (M+1)st row will be consistently estimated as a function of the first row of A.

Define the solution to the iteration on (31) as the (set) function

$$(\gamma, \delta) = \phi(\alpha, \beta) \tag{32}$$

ϕ maps the α's and β's into a set of γ's and δ's that satisfy restriction (24). Hence, one (consistent) estimator of γ, δ is $\phi(\alpha, \beta)$.

Under the restriction (24) the likelihood function in (25), $L(\alpha, \beta, \gamma, \delta|\{s_t\}, \{f_t\})$, becomes a function only of the α's and β's. As Wilson (1973) argues, maximum likelihood estimtes with an unknown V are obtained by minimizing $|\hat{V}|$, with respect to the α's and β's, where

$$|\hat{V}| = |\frac{1}{T}\sum_{t=1}^{T} e_t(\alpha, \beta)\, e_t(\alpha, \beta)'|$$

where the $e_t(\alpha, \beta)$, the residuals from (14), are functions of the α's and β's only, since they were calculated from (14) with (24) imposed.

A derivative free nonlinear minimization routine can be used to estimate (14) under the restriction (24). The IMSL subroutine ZXMIN, which uses a quasi-Newton method, was used. Generally, 600 iterations were required to obtain three significant digits. The least squares estimates of α and β were used as starting values.

Tables 1 and 2 report, in addition to the unrestricted estimtes, the restricted maximum likelihood estimates. In addition, the likelihood ratio statistic, which is distributed $\chi^2(8)$, is reported along with the marginal significance level.

The results for Germany indicate a failure to reject the null hypothesis. The likelihood ratio statistic of 11.134 is insignificant, as indicated by the marginal significance level of 19.4 percent. The results for Canada, a likelihood ratio statistic of 20.843 (marginal significance level of 0.8 percent), indicate a rejection of the null hypothesis.

Not surprisingly, the results implied by the W-statistic and likelihood ratio statistic are very close (marginal significance levels of 23.7 percent and 19.4 percent for Germany, and 0.5 percent and 0.8 percent for Canada). This is as it should be, since both tests are asymptotically equivalent (Silvey 1970, p. 118), and the sample size is 188 observations.

Can we use the restricted estimates to locate the reason the null hypothesis is rejected for Canada? Table 3 presents an estimate of the variance-covariance matrix of the restricted parameter estimates (α and β) for Canada. Standard errors for γ and δ are not reported since γ and δ cannot be analytically solved for in terms of α and β. The difference between the restricted and unrestricted estimates of α and β are insignificant. Using the standard errors of the unrestricted estimates of γ and δ (not reported), the unrestricted estimate of γ_1 is significantly different from the restricted estimate (all other γ's and δ's are insignificantly different). In calculating the Wald statistic we required an estimate of $h(\psi)$--the restriction vector implied by (24). An estimate of that vector, evaluated at the OLS estimate is

TABLE 3

VARIANCE-COVARIANCE MATRIX OF CANADIAN ESTIMATES OF (α, β)

	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
α_1	0.171							
α_2	-0.133	0.605						
α_3	0.007	0.331	0.386					
α_4	0.243	-0.247	0.004	0.523				
β_1	0.035	0.036	-0.005	0.027	0.030			
β_2	-0.228	-0.179	-0.107	-0.221	-0.133	0.683		
β_3	-0.080	0.066	-0.333	-0.116	0.083	-0.072	0.755	
β_4	0.074	-0.115	-0.070	-0.257	0.028	-0.159	-0.115	0.848

$$h = (0.211 \quad 0.129 \quad 0.079 \quad 0.035 \quad -0.185 \quad -0.052 \quad -0.032 \quad -0.00)$$

$$|h - 0| = 0.107.$$

It is interesting to note that the first restriction, $h_1(\psi) = 0.211$, is also the largest in absolute value. Again the results are consistent between the Wald test and likelihood ratio test.

The period being examined, April 1973 - May 1975, was a time of adjustment to a new floating rate system. It is possible that the "noisiness" of the system changed over time (Hakkio [1979] presents evidence that the variance had increased for Canada). To allow for this possibility, we split the Canadian sample in two, and reestimated the restricted and unrestricted version of the model. The results are presented in Tables 4 and 5.

For the first period we obtain a likelihood ratio statistic 21.465, with a marginal significance level of 0.6 percent. So, we again reject the null hypothesis for Canada for the first period. The likelihood ratio statistic for the whole period and the first period are quite close. For the second period we obtain a likelihood ratio statistic of only 13.302, with a marginal significance level of 10.2 percent. Therefore, for the second period we fail to reject the null hypothesis, and therefore conclude that the data are consistent with the pure expectations theory of the term structure of the forward premium--for the second period. As before, the roots of the A matrix indicate that the process was stationary. As shown in Hakkio (1979), the variance of the forecast error changed over time in the Canadian case. Consequently, the data may not have been stationary and so the econometric methodology may have been suspect. Splitting the sample in two presumably reduced the problem posed by heteroskedasticity.

VI. Conclusions

Most studies that test foreign exchange market efficiency focus on the relation between the spot exchange rate and a single maturity for the forward exchange rate, usually the one month rate. This procedure ignores the fact that more than one maturity currently is being traded. This paper extends the analysis of market efficiency so as to obtain implications concerning the joint movements of the spot exchange rate and the one and six month forward exchange rate. Using the certainty equivalence theory of the term structure of (domestic and foreign) interest rates and the

TABLE 4

CANADA: ESTIMATES OF BIVARIATE AUTOREGRESSION
UNRESTRICTED AND RESTRICTED, FIRST PERIOD

j	1	2	3	4
		Unrestricted Estimates		
α_j	-0.5758	-0.2897	-0.3780	-0.4209
β_j	1.0329	0.4575	0.7801	0.0494
γ_j	0.0177	-0.0767	-0.1543	-0.1640
δ_j	0.2066	0.2860	0.3201	0.0760

$$\hat{V} = \begin{bmatrix} 11.125*10^{-8} & 4.6840*10^{-8} \\ & 2.8978*10^{-8} \end{bmatrix} \qquad |\hat{V}| = 1.030*10^{-15}$$

j	1	2	3	4
		Restricted Estimates		
α_j	-0.7581	-0.2875	-0.1967	-0.2003
β_j	0.9032	0.1384	0.3603	-0.0012
γ_j	-0.0895	-0.0462	-0.0227	-0.0098
δ_j	0.0880	0.0302	0.0176	-0.0001

$$\hat{V} = \begin{bmatrix} 12.691*10^{-8} & 5.748*10^{-8} \\ & 3.631*10^{-9} \end{bmatrix} \qquad |\hat{V}| = 1.304*10^{-15}$$

Likelihood ratio statistic = 21.465

Marginal significance level = 0.006

TABLE 5

CANADA: ESTIMATES OF BIVARIATE AUTOREGRESSION
UNRESTRICTED AND RESTRICTED, SECOND PERIOD

j	1	2	3	4
		<u>Unrestricted Estimates</u>		
α_j	-0.0554	0.0691	0.1566	0.2611
β_j	0.0727	-0.0466	-0.1590	-0.1747
γ_j	0.3164	0.2144	0.1989	0.3244
δ_j	-0.2859	-0.1703	-0.1966	-0.1570

$$\hat{V} = \begin{pmatrix} 5.932*10^{-8} & 3.503*10^{-8} \\ & 4.736*10^{-8} \end{pmatrix} \qquad |\hat{V}| = 1.582*10^{-15}$$

j	1	2	3	4
		<u>Restricted Estimates</u>		
α_j	-0.1480	0.1212	0.2087	0.2047
β_j	0.1487	-0.0993	-0.1770	-0.1684
γ_j	0.0518	0.0964	0.0822	0.0437
δ_j	-0.0376	-0.0802	-0.0688	-0.0360

$$\hat{V} = \begin{pmatrix} 6.013*10^{-8} & 3.649*10^{-8} \\ & 5.259*10^{-8} \end{pmatrix} \qquad |\hat{V}| = 1.831*10^{-15}$$

Likelihood ratio statistic = 13.302

Marginal significance level = 0.102

hypothesis of interest rate parity, it is possible to write the six month forward premium as a geometric average of the current one month forward premium and expected future one month forward premiums. Testable implications are then derived by assuming that the one and six month forward premiums are generated by a bivariate autoregression.

The hypothesis of rational expectations imposes a set of highly non-linear cross equation restrictions on the parameters of the model. Two different methods were then presented to test the validity of the restrictions. Both methods yielded identical conclusions. It was found that for Germany the data were consistent with the theory of the term structure of the forward premium. For Canada the data were inconsistent with the theory. Hakkio (1979) found that the Canadian - U.S. exchange rate system appeared to change over time with respect to the forecast variance. To examine the implications of this, we estimated the model over two subperiods. We then found that although the data were inconsistent with the theory in the first subperiod, the data were consistent with the theory in the second subperiod. Due to the complex nature of the restrictions, the cause of rejection could not be fully ascertained. In addition, although the theory may be rejected for Canada, there is no clear alternative theory that can be accepted.

REFERENCES

Anderson, T. W. The Statistical Analysis of Time Series. New York: John Wiley and Sons, Inc., 1971.

Box, George E. P. and Jenkins, Gwilym M. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 1976.

Brillembourg, Arturo. "The Term Structure of Forward Rates." Unpublished manuscript, International Monetary Fund, Washington, D.C., 1978.

Dooley, Michael P. and Shafer, Jeffrey R. "Analysis of Short-run Exchange rate Behavior: March 1973 to September 1975." International Finance Discussion Paper, No. 76. Federal Reserve System, Washington, D.C., February 1976.

Fama, Eugene. "Efficient Capital Markets: A Review of Theory and Empirical Work." In Frontiers of Quantitative Economics. Edited by Michael Intriligator. Amsterdam: North Holland Publishing Company, 1971.

Frenkel, Jacob A. and Levich, Richard M. "Transaction Costs and Interest Arbitrage: Tranquil versus Turbulent Periods." Journal of Political Economy 85 (December 1977): 1209-1226.

__________. "Covered Interest Arbitrage: Unexploited Profits?" Journal of Political Economy 83 (April 1975): 325-338.

Giddy, Ian H. "Term Structure and Expectations in the Money and Foreign Exchange Market." Unpublished manuscript, University of Chicago, March 1977.

Hakkio, Craig S. "Expectations and the Foreign Exchange Market." Ph.D. dissertation, University of Chicago, 1979.

Hansen, Lars Peter. "The Asymptotic Distribution of Least Squares Estimators with Endogenous Regressors and Dependent Residuals." Unpublished manuscript, Carnegie-Mellon University, March 1979.

Jensen, Michael C. "Some Anomalous Evidence Regarding Market Efficiency." Journal of Financial Economics 6 (January 1978): 95-101.

Kessel, Reuben. The Cyclical Behavior of the Term Structure of Interest Rates. National Bureau of Economic Research Occasional Paper 91. New York: Columbia University Press.

Keynes, John M. A Tract on Monetary Reform. London: Macmillan and Co., Ltd., 1923.

Kohlhagen, Steven W. "The Foreign Exchange Markets--Models, Tests and Empirical Evidence." Unpublished manuscript, University of California, Berkeley, 1976.

Koopmans, L. R. The Spectral Analysis of Time Series. New York: Academic Press, 1974.

Kouri, Pentti. "The Determinants of the Forward Premium." Center for Research in Economic Growth, Memorandum No. 204, Stanford University, July 1976.

Levich, Richard M. "On the Efficiency of Markets for Foreign Exchange." In International Economic Policy. Edited by Rudiger Dornbusch and Jacob A. Frenkel. Baltimore: The Johns Hopkins University Press, 1979.

Meiselman, David. The Term Structure of Interest Rates. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

Modigliani, Franco and Shiller, Robert J. "Inflation, Rational Expectations and the Term Structure of Interest Rates." Economics, N.S. 40 (February 1973): 12-43.

Morrison, Donald F. Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1967.

Muth, John F. "Rational Expectations and the Theory of Price Movements." Econometrica 29 (July 1961): 315-335.

Nelson, Charles. The Term Structure of Interest Rates. New York: Basic Books, Inc., 1972.

Porter, Michael G. "A Theoretical and Empirical Framework for Analyzing the Term Structure of Exchange Rate Expectations." IMF Staff Papers 18 (November 1971): 613-645.

Rao, C. R. Linear Statistical Inference and its Applications. 2nd ed. New York: John Wiley and Sons, 1973.

Roll, Richard. "The Efficient Market Model Applied to the U.S. Treasury Bill Rates." Ph.D. dissertation, University of Chicago, 1968.

Sargent, Thomas J. "Rational Expectations and the Term Structure of Interest Rates." Journal of Money, Credit, and Banking 4 (February 1972): 74-97.

__________. Macroeconomic Theory. New York: Academic Press, 1979. (a)

__________. "A Note on Maximum Likelihood Estimation of the Rational Expectation Model of the Term Structure." Journal of Monetary Economics 5 (January 1979): 133-143. (b)

Silvey, S. D. Statistical Inference. New York: Chapman Hall, 1975.

Telser, Lester G. "A Critique of Some Recent Empirical Research on the Explanation of the Term Structure of Interest Rates." Journal of Political Economy 75 (August 1967): 546-560.

Whittle, Paul. Prediction and Regulation by Linear Least Square Methods. Princeton, N.J.: D. Van Nostrand Company, Inc., 1963.

Wilson, G. Tunnicliffe. "The Estimation of Parameters in Multivariate Time Series Models." Journal of the Royal Statistical Society B 35, No. 1 (1973): 76-85.