Cointegration and Forward and Spot Exchange Rate Regressions

by

Eric Zivot*
Department of Economics
University of Washington
Box 353330
Seattle, WA 98195-3330
ezivot@u.washington.edu

June 12, 1997
Last Revision: September 29, 1998

Abstract

In this paper we investigate in detail the relationship between models of cointegration between the current spot exchange rate, s_t, and the current forward rate, f_t, and models of cointegration between the future spot rate, s_{t+1}, and f_t and the implications of this relationship for tests of the forward rate unbiasedness hypothesis (FRUH). We argue that simple models of cointegration between s_t and f_t more easily capture the stylized facts of typical exchange rate data than simple models of cointegration between s_{t+1} and f_t and so serve as a natural starting point for the analysis of exchange rate behavior. We show that simple models of cointegration between s_t and f_t imply rather complicated models of cointegration between s_{t+1} and f_t. As a result, standard methods are often not appropriate for modeling the cointegrated behavior of $(s_{t+1}, f_t)'$ and we show that the use of such methods can lead to erroneous inferences regarding the FRUH.

*Thanks to Charles Engel for his insights on the issues presented herein and to Vinay Datar for motivating me to write the paper. Updates to the paper can be found at http://weber.u.washington.edu/~ezivot/homepage.htm.

1. Introduction

There is an enormous literature on testing if the forward exchange rate is an unbiased predictor of future spot exchange rates. Engel (1996) provides the most recent review. The earliest studies, e.g. Cornell (1977), Levich (1979) and Frenkel (1980), were based on the regression of the log of the future spot rate, s_{t+1}, on the log of the current forward rate, f_t. The results of these studies generally support the forward rate unbiasedness hypothesis (FRUH). Due to the unit root behavior of exchange rates and the concern about the spurious regression phenomenon illustrated by Granger and Newbold (1974), later studies, e.g. Bilson (1981), Fama (1984) and Froot and Frankel (1989), concentrated on the regression of the change in the log spot rate, Δs_{t+1}, on the forward premium, $f_t - s_t$. Overwhelmingly, the results of these studies reject the FRUH. The most recent studies, e.g., Hakkio and Rush (1989), Barnhart and Szakmary (1991), Naka and Whitney (1995), Hai, Mark and Yu (1997), Norrbin and Reffett (1996), Newbold, et. al. (1996), Clarida and Taylor (1997), Barnhart, McNown and Wallace (1998) and Luintel and Paudyal (1998) have focused on the relationship between cointegration and tests of the FRUH. The results of these studies are mixed and depend on how cointegration is modeled.

Since the results of Hakkio and Rush (1989), it is well recognized that the FRUH requires that s_{t+1} and f_t be cointegrated and that the cointegrating vector be (1,-1) and much of the recent literature has utilized models of cointegration between s_{t+1} and f_t. It is also true that the FRUH requires s_t and f_t to be cointegrated with cointegrating vector (1,-1) and only a few authors have based their analysis on models of cointegration between s_t and f_t . In this paper we investigate in detail the relationship between models of cointegration between s_t and f_t and models of cointegration between s_{t+1} and f_t and the implications of this relationship for tests of the FRUH. We argue that simple models of cointegration between s_t and f_t more easily capture the stylized facts of typical exchange rate data than simple models of cointegration between s_{t+1} and f_t and so serve as a natural starting point for the analysis of exchange rate behavior. Simple models of cointegration between s_t and f_t imply rather complicated models of cointegration between s_{t+1} and f_t. In particular, starting with a first order bivariate vector error correction model for $(s_t, f_t)'$ we show that the implied cointegrated model for $(s_{t+1}, f_t)'$ is nonstandard and does not have a finite VAR representation. As a result, standard VAR methods are not appropriate for modeling $(s_{t+1}, f_t)'$ and we show that the use of such methods can lead to erroneous inferences regarding the FRUH. In particular, we show that tests of the null of no-cointegration based on common cointegrated models for $(s_{t+1}, f_t)'$ are likely to be severely size distorted. In addition, using the implied triangular cointegrated representation for $(s_{t+1}, f_t)'$ we can explicitly characterize the OLS bias in the levels regression of s_{t+1} on f_t. Based on this representation, we can show that the OLS estimate of the coefficient on f_t in the levels regression is downward biased (away from one) even if the FRUH is true. Finally, we show that the results of Naka and Whitney (1995) and Norrbin and Reffett (1996) supporting the FRUH are driven by the specification of the cointegration model for $(s_{t+1}, f_t)'$.

The plan of the paper is as follows. In section 2, we discuss the relationship between cointegration and the tests of the forward rate unbiasedness hypothesis. In section 3 we present some stylized facts of exchange rate data typically used in investigations of the FRUH. In section 4, we discuss some simple models of cointegration between s_t and f_t that capture the basic stylized facts about the data and we show the restrictions that the FRUH places on these models. In section 5, we consider models of cointegration between s_{t+1} and f_t that are implied by models of cointegration between s_t and f_t. In section 6, we use our results to reinterpret some recent findings concerning the FRUH reported by Naka and Whitney (1995) and Norrbin and Reffett (1996). Our concluding remarks are given in section 7.

2. Cointegration and the Forward Rate Unbiasedness Hypothesis: An Overview

The relationship between cointegration and the forward rate unbiasedness hypothesis has been discussed by several authors starting with Hakkio and Rush (1987). Engel (1996) provides a comprehensive review of this literature and serves as a starting point for the analysis in this paper. Following Engel (1996), the forward exchange rate unbiasedness hypothesis (FRUH) *under rational expectations and risk neutrality* is given by

$$E_t[s_{t+1}] = f_t, \tag{1}$$

where $E_t[\cdot]$ denotes expectation conditional on information available at time t. Using the terminology of Baillie (1989), FRUH is an example of the "observable expectations" hypothesis. The FRUH is usually expressed as the levels relationship

$$s_{t+1} = f_t + \xi_{t+1}, \tag{2}$$

where ξ_{t+1} is a random variable (rational expectations forecast error) with $E_t[\xi_{t+1}] = 0$. It should be kept in mind that rejection of the FRUH can be interpreted as a rejection of the model underlying $E_t[\cdot]$ or a rejection of the equality in (1) itself.

Two different regression equations have generally been used to test the FRUH. The first is the "levels regression"

$$s_{t+1} = \mu + \beta_f f_t + u_{t+1} \tag{3}$$

and the null hypothesis that FRUH is true imposes the restrictions $\mu = 0$, $\beta_f = 1$ and $E_t[u_{t+1}] = 0$. In early empirical applications authors generally focused on testing the first two conditions and either ignored the latter condition or informally tested it using the Durbin-Watson statistic. Most studies using (3) found estimates of β_f very close to 1 and hence supported the FRUH. Some authors, e.g. Barnhart and Szakmary (1991), Liu and Maddala (1992), Naka and Whitney (1995) and Hai, Mark and Wu (1997), refer to testing $\mu = 0$, $\beta_f = 1$ as testing the forward rate unbiasedness condition (FRUC). Testing the orthogonality condition $E_t[u_{t+1}] = 0$, conditional on not rejecting FRUC, is then referred to as testing forward market efficiency under rational expectaions and risk neutrality. Assuming s_t and f_t have unit roots, i.e., $s_t, f_t \sim I(1)$, (see, for example, Messe and Singleton (1982), Baillie and Bollerslev (1989), Mark (1990), Liu and Maddala (1992), Crowder (1994), or Clarida and Taylor (1997) for empirical evidence), then the FRUH requires that s_{t+1} and f_t be cointegrated with cointegrating vector (1, -1) *and* that the stationary, i.e., $I(0)$, cointegrating residual, u_{t+1}, satisfy $E_t[u_{t+1}] = 0$. Notice that testing FRUC is then equivalent to testing for cointegration between s_{t+1} and f_t and that the cointegrating vector is (1,-1) and testing forward market efficiency is equivalent to testing that the forecast error, $s_{t+1} - f_t$, has conditional mean zero.

The FRUH assumes rational expectations and risk neutrality. Under rational expectations, if agents are risk averse then a stationary time-varying risk premium exists and the relationship between s_{t+1} and f_t becomes

$$s_{t+1} = f_t - rp_t^{re} + \xi_{t+1}, \tag{4}$$

where $rp_t^{re} = f_t - E_t[s_{t+1}]$ represents the stationary rational expectations risk premium. As long as rp_t^{re} is stationary s_{t+1} and f_t will be cointegrated with cointegrating vector (1, -1) but f_t will be a biased predictor of s_{t+1} provided rp_t^{re} is predictable using information available at time t. In this regard, the

FRUC is a misleading acronym since if the FRUC is true forward rates are not necessarily unbiased predictors of future spot rates. Since the FRUC is equivalent to cointegration between s_{t+1} and f_t and that the cointegrating vector is (1,-1), which implies that s_{t+1} and f_t trend together in the long-run but may deviate in the short-run, it is more appropriate to call this the *long-run* forward rate unbiasedness condition (LRFRUC).

Several authors, e.g. Messe and Singleton (1982), Meese (1989) and Isard (1995), have stated that since s_t and f_t have unit roots the levels regression (3) is not a valid regression equation because of the spurious regression problem described in Granger and Newbold (1974). However, this is not true if s_{t+1} and f_t are cointegrated. What is true is that if s_{t+1} and f_t are cointegrated with cointegrating vector (1, -1), which allows for the possibility of a stationary time varying risk premium so that u_{t+1} in (3) is $I(0)$, then the OLS estimates from (3) will be super consistent (converge at rate T instead of rate $T^{1/2}$) for the true value $\beta_f = 1$ but generally not efficient and biased away from 1 in finite samples so that the asymptotic distributions of t-tests and F-tests on μ an β_f will follow non-standard distributions, see Corbae, Lim and Ouliaris (1992). Hence, even if the FRUH is not true due the existence of a stationary risk premium so that (3) is a misspecified regression, OLS on (3) still gives a consistent estimate of $\beta_f = 1$. More importantly, there are simple modifications to OLS that yield asymptotically unbiased and efficient estimates of the parameters of (3) in the presence of general serial correlation and heteroskedasticity and these modifications should be used to make inferences about the parameters in the levels regression[1]. In this sense, the levels regression (3) under cointegration is asymptotically immune to the omission of a stationary risk premium. Hai, Mark and Wu (1997) use Stock and Watson's (1993) dynamic OLS estimator on the levels regression (3) and provide evidence that s_{t+1} and f_t are cointegrated with cointegrating vector (1, -1).

The second regression equation used to test the FRUH is the "differences equation"

$$\Delta s_{t+1} = \mu^* + \alpha_s(f_t - s_t) + u^*_{t+1} \tag{5}$$

and the null hypothesis that FRUH is true imposes the restrictions $\mu^* = 0$, $\alpha_s = 1$ and $E_t[u^*_{t+1}] = 0$. Empirical results based on (5), surveyed in Engel (1996), overwhelmingly reject the FRUH. In fact, typical estimates of α_s across a wide range of currencies and sampling frequencies are significantly negative. This result is often referred to as the forward discount anomaly, forward discount bias or forward discount puzzle and seems to contradict the results based on the levels regression (3). Given that $s_t, f_t \sim I(1)$, for (5) to be a "balanced regression" (i.e., all variables in the regression are integrated of the same order) the forward premium, $f_t - s_t$, must be $I(0)$ or, equivalently, f_t and s_t must be cointegrated with cointegrating vector (1,-1). Assuming that covered interest rate parity holds, the forward premium is simply the interest rate differential between the respective countries and there are good economic reasons to believe that such differentials do not contain a unit root. Hence, tests of the FRUH based on (5) implicitly assume that the forward premium is $I(0)$ and so such tests are conditional on f_t and s_t being cointegrated with cointegrating vector (1,-1). In this respect, (5) can be thought of as one equation in a particular vector error correction model (VECM) for $(f_t, s_t)'$.[2] Horvath and Watson (1995) and Clarida and Taylor (1997) use VECM-based tests and provide evidence that f_t and s_t are cointegrated with cointegrating vector (1,-1).

As noted by Fama (1984), the negative estimates of α_s are consistent with rational expectations and market efficiency and imply certain restrictions on the risk premium. To see this, note that under rational expectations we may write

$$\Delta s_{t+1} = (f_t - s_t) - rp_t^{re} + \xi_{t+1}, \tag{6}$$

so that the difference regression (5) is misspecified if risk neutrality fails. Since all variables in (6) are $I(0)$, if the risk premium is correlated with the forward premium then the OLS estimate of α_s in the standard differences regression (5), which omits the risk premium, will be biased away from the true value of 1. Hence the negative estimates of α_s from (5) can be interpreted as resulting from omitted variables bias. As discussed in Fama (1984), for omitted variables bias to account for negative estimates of α_s it must be true that $cov(E_t[s_{t+1}] - s_t, rp_t^{re}) < 0$ and $var(rp_t^{re}) > Var(E_t[s_{t+1}] - s_t)$. Hence, as Engel (1996) notes, models of the foreign exchange risk premium should be consistent with these two inequalities.

The tests of the FRUH based on (3) and (5) involve cointegration either between s_{t+1} and f_t or f_t and s_t. As Engel (1996) points out, since

$$s_{t+1} - f_t = \Delta s_{t+1} - (f_t - s_t)$$

it is trivial to see under the assumption that f_t and s_t are $I(1)$ that (i) if s_t and f_t are cointegrated with cointegrating vector (1,-1) then s_{t+1} and f_t must be cointegrated with cointegrating vector (1,-1); and (ii) if s_{t+1} and f_t are cointegrated with cointegrating vector (1,-1) then s_t and f_t must be cointegrated with cointegrating vector (1,-1). Cointegration models for $(s_t, f_t)'$ and $(s_{t+1}, f_t)'$ can both be used to describe the data and test the FRUH but the form of the models used can have a profound impact on the resulting inferences. For example, we show that a simple first order vector error correction model for $(s_t, f_t)'$ describes monthly data well and leads naturally to the differences regression (5) from which the FRUH is easily rejected. In contrast, we show that some simple first order vector error correction models for $(s_{t+1}, f_t)'$, which are used in the empirical studies of Naka and Whitney (1995) and Norrbin and Reffett (1996), miss some important dynamics in monthly data and as a result indicate that the FRUH appears to hold. Hence, misspecification of the cointegration model for $(s_{t+1}, f_t)'$ can explain some of the puzzling empirical results concerning tests of the FRUH.

In the next section, we describe some stylized facts of monthly exchange rate data that are typical in the analysis of the FRUH. In the remaining sections we use these facts to motivate certain models of cointegration for $(s_t, f_t)'$ and $(s_{t+1}, f_t)'$ to support our claims regarding misspecification and tests of the FRUH.

3. Some Stylized Facts of Typical Exchange Rate Data

Let f_t denote the log of the forward exchange rate in month t and s_t denote the log of the spot exchange rate. We focus on monthly data for which the maturity date of the forward contract is the same as the sampling interval to avoid modeling complications created by overlapping data. For our empirical examples, we consider forward and spot rate data (all relative to the U.S. dollar) on the pound, yen and Canadian dollar taken from Datastream[3]. Figure 1 shows time plots of s_{t+1}, f_t - s_t (forward premium), and s_{t+1} - f_t (forecast error) for the three currencies and Table 1 gives some summary statistics of the data. Spot and forward rates behave very similarly and exhibit random walk type behavior. The forward premiums are all highly autocorrelated but the forecast errors show very little autocorrelation. The variances of Δs_{t+1} and Δf_{t+1} are roughly ten times larger than the variance of f_t - s_t and are similar to the variance of s_{t+1} - f_t. For all currencies, Δs_{t+1}, Δf_{t+1} and s_{t+1} - f_t are negatively correlated with f_t - s_t. Any model of cointegration with cointegrating vector (1, -1) for $(s_t, f_t)'$ or $(s_{t+1}, f_t)'$ should capture these basic stylized facts.

4. Models of Cointegration between f_t and s_t

4.1 Vector error correction representation

The stylized facts of the monthly exchange rate data reported in the previous section can be captured by a simple cointegrated VAR(1) model for $y_t = (f_t, s_t)'$. This simple model has also recently been used by Godbout and van Norden (1996). Before presenting the empirical restults, we begin this section with a review of the properties of such a model. The general bivariate VAR(1) model for y_t is

$$y_t = \mu + \Phi y_{t-1} + \epsilon_t,$$

where $\epsilon_t \sim iid\ (0,\Sigma)$ and Σ has elements σ_{ij} ($i,j = f,s$), and can be reparameterized as

$$\Delta y_t = \mu + \Pi y_{t-1} + \epsilon_t \tag{7}$$

where $\Pi = \Phi - I$. Under the assumption of cointegration, Π has rank 1 and there exist 2×1 vectors α and β such that $\Pi = \alpha\beta'$. Using the normalization $\beta = (1, -\beta_s)'$, (7) becomes the vector error correction model (VECM)

$$\Delta f_t = \mu_f + \alpha_f(f_{t-1} - \beta_s s_{t-1}) + \epsilon_{ft}, \tag{8a}$$

$$\Delta s_t = \mu_s + \alpha_s(f_{t-1} - \beta_s s_{t-1}) + \epsilon_{st}. \tag{8b}$$

Since spot and forward rates often do not exhibit a systematic tendency to drift up or down it may be more appropriate to restrict the intercepts in (8) to the error correction term. That is, $\mu_f = -\alpha_f \mu_c$ and $\mu_s = -\alpha_s \mu_c$. Under this restriction s_t and f_t are $I(1)$ without drift and the cointegrating residual, $f_t - \beta_s s_t$, is allowed to have a nonzero mean μ_c[4].

With the intercepts in (8) restricted to the error correction term, the VECM can be solved to give a simple AR(1) model for the cointegrating residual $\beta' y_t - \mu_c = f_t - \beta_s s_t - \mu_c$. Premultiplying (7) by β' and rearranging gives

$$f_t - \beta_s s_t - \mu_c = \phi(f_{t-1} - \beta_s s_{t-1} - \mu_c) + \eta_t, \tag{9}$$

where $\phi = 1 + \beta'\alpha = 1 + (\alpha_f - \beta_s\alpha_s)$ and $\eta_t = \beta'\epsilon_t = \epsilon_{ft} - \beta_s\epsilon_{st}$. Since (9) is simply an AR(1) model, the cointegrating residual is stable and stationary if $|\phi| = |1 + (\alpha_f - \beta_s\alpha_s)| < 1$. Notice that if $\alpha_f = \beta_s\alpha_s$ then the cointegrating residual is $I(1)$ and f_t and s_t are not cointegrated.

The exogeneity status of spot and forward rates with regard to the cointegrating parameters α and β was the focus of attention in Norrbin and Reffett (1996) so it is appropriate to discuss this issue in some detail. Exogeneity issues in error correction models are discussed at length in Johansen (1992, 1995), Banerjee *et. al.* (1993), Urbain (1993), Ericsson and Irons (1994) and Zivot (1998). For our purposes weak exogeneity of spot or forward rates for the cointegrating parameters α and

β places restrictions on the parameters of the VECM (8). In particular, if f_t is weakly exogenous with respect to $(\alpha_s, \beta_s)'$ then $\alpha_f = 0$ and efficient estimation of the cointegrating parameters can be made from the single equation conditional error correction model

$$\Delta s_t = \mu_s + \alpha_s(f_{t-1} - \beta_s s_{t-1}) + \gamma_s \Delta f_t + \nu_{st}, \tag{10a}$$

where $\gamma_s = \sigma_{ss}^{-1}\, \sigma_{fs}$ and ν_{st} is uncorrelated with ϵ_{ft}. Similarly, if s_t is weakly exogenous with respect to $(\alpha_f, \beta_s)'$ then $\alpha_s = 0$ and efficient estimation of the cointegrating parameters can be made from the single equation conditional error correction model

$$\Delta f_t = \mu_f + \alpha_f(f_{t-1} - \beta_s s_{t-1}) + \gamma_f \Delta s_t + \nu_{ft} \tag{10b}$$

where $\gamma_f = \sigma_{ff}^{-1}\, \sigma_{fs}$ and ν_{ft} is uncorrelated with ϵ_{st}.

If $\beta_s = 1$ then the forward premium is $I(0)$ and follows an AR(1) process and the VECM (8) becomes

$$\Delta f_t = \mu_f + \alpha_f(f_{t-1} - s_{t-1}) + \epsilon_{ft}, \tag{11a}$$

$$\Delta s_t = \mu_s + \alpha_s(f_{t-1} - s_{t-1}) + \epsilon_{st}. \tag{11b}$$

Notice that (11b) is simply the standard differences regression (5) used to test the FRUH. Further, if α_f and α_s are of the same sign and magnitude then the implied value of ϕ in (9) is close to 1 and this corresponds to the stylized fact that the forward premium is stationary but very highly autocorrelated. Also, the implied variance of the of the forward premium from (9) is $\sigma_{\eta\eta} = \sigma_{ff} + \sigma_{ss} - 2\rho_{fs}(\sigma_{ff}\sigma_{ss})^{1/2}$ and will be very small relative to the variances of Δf_t and Δs_t given the stylized facts that $\sigma_{ff} \approx \sigma_{ss}$ and $\rho_{fs} \approx 1$.

The FRUH places testable restrictions on the VECM (8). Necessary conditions for the FRUH to hold are (i) s_t and f_t are cointegrated (ii) $\beta_s = 1$ and (iii) $\mu_c = 0$. In addition, the FRUH requires that $\alpha_s = 1$ in order for the forecast error in (2) to have conditional mean zero. It is important to stress that, together, these two restrictions limit both the long-run *and* short-run behavior of spot and forward rates. Applying these restrictions, (8), led one period, becomes

$$\Delta f_{t+1} = \alpha_f(f_t - s_t) + \epsilon_{f,t+1}, \tag{12a}$$

$$\Delta s_{t+1} = (f_t - s_t) + \epsilon_{s,t+1}. \tag{12b}$$

Notice that the FRUH requires that the expected change spot rate is equal to the forward premium or, equivalently, that the adjustment to long-run equilibrium occurs in one period. The change in the forward rate, on the other hand, is directly related to the persistence of interest rate differentials now measured by α_f since $\phi = 1 + (\alpha_f - 1) = \alpha_f$. Stability of the VECM under the FRUH requires that $|\alpha_f| < 1$. Thus, the FRUH is consistent with a highly persistent forward premium.

The representation in (12) shows that weak exogeneity of spot rates with respect to the cointegrating parameters is inconsistent with the FRUH because if spot rates are weakly exogenous then $\alpha_s = 0$ and the FRUH cannot hold. In addition if the FRUH is true and forward rates are weakly exogenous then (12) cannot capture the dynamics of typical data. To see this, suppose that forward rates are weakly exogenous so that $\alpha_f = 0$. Since $\sigma_{ss} \approx \sigma_{ff} \approx \sigma_{sf} = \sigma^2$ it follows that $\gamma_s \approx \gamma_f \approx 1$. If $\mu_s = 0$, $\alpha_s = 1$ and $\beta_s = 1$, then (10a) becomes

$$s_t = f_{t-1} + \Delta f_t + \nu_t = f_t + \nu_t$$

which simply states that the current spot rate is equal to the current forward rate plus a white noise error. This result is clearly inconsistent with the data since it implies that the forward premium is serially uncorrelated.

4.2 Phillips' triangular representation.

Another useful representation of a cointegrated system is Phillips' (1991) triangular representation, which is similar to the triangular representation of a limited information simultaneous equations model. This representation is most useful for studying the asymptotic properties of cointegrating regressions and Baillie (1989) has advocated its use for testing rational expectations restrictions in cointegrated VAR models. This representation is also used by Naka and Whitney (1995) to test the FRUH. The general form of the triangular representation for y_t is

$$f_t = \mu_c + \beta_s s_t + u_{ft}, \tag{13a}$$

$$s_t = s_{t-1} + u_{st}, \tag{13b}$$

where the vector of errors $u_t = (u_{ft}, u_{st})' = (f_t - \beta_s s_t - \mu_c, \Delta s_t)'$ has the stationary moving average representation $u_t = \psi(L)e_t$ where $\psi(L) = \sum_{k=0}^{\infty} \psi_k L^k$, $\sum_{k=0}^{\infty} |k|\psi_k < \infty$ and e_t is *i.i.d.* with mean zero and covariance matrix V. Equation (13a) models the (structural) cointegrating relationship and (13b) is a reduced form relationship describing the stochastic trend in the spot rate. For a given VECM representation, the triangular representation is simply a reparameterization. For the VECM (8) with the restricted constant, the derived triangular representation is given by (13a)-(13b) with

$$u_{ft} = \phi u_{f,t-1} + \eta_t, \tag{13c}$$

$$u_{st} = \alpha_s u_{ft-1} + \epsilon_{st}, \qquad (13d)$$

and ϕ, α_s, η_t and ϵ_{st} are as previously defined. Equation (13c) models the disequilibrium error (which equals the forward premium if $\beta_s = 1$) as an AR(1) process and (13d) allows the lagged error to affect the change in the spot rate. Let $e_t = (\eta_t, \epsilon_{st})'$. Then the vector $u_t = (u_{ft}, u_{st})' = (f_t - \beta_s s_t - \mu_c, \Delta s_t)'$ has the VAR(1) representation $u_t = Cu_{t-1} + e_t$ where

$$C = \begin{pmatrix} \phi & 0 \\ \alpha_s & 0 \end{pmatrix}, \; V = \begin{pmatrix} \sigma_{\eta\eta} & \sigma_{\eta s} \\ \sigma_{s\eta} & \sigma_{ss} \end{pmatrix},$$

Hence, $\psi(L) = (I - CL)^{-1}$. As noted previously, $\sigma_{\eta\eta}$ is very small relative to σ_{ss} and $\sigma_{\eta s} = \sigma_{fs} - \sigma_{ss}$.

Phillips and Loretan (1991) and Phillips (1991) show how the triangular representation of a cointegrated system can be used to derive the asymptotic properties of the OLS estimates of the cointegration parameters. For our purposes, the most important result is that the OLS estimate of β_s from (13a) is asymptotically unbiased and efficient only if u_{ft} and u_{st} are contemporaneously uncorrelated and there is no feedback between u_{ft} and u_{st} (i.e, s_t is weakly exogenous for β_s and u_{ft} does not Granger cause u_{st} and vise-versa). These correlation and feedback effects can be expressed in terms of specific components of the long-run covariance matrix of u_t. The long-run covariance matrix of u_t is defined as $\Omega = \sum_{k=-\infty}^{\infty} E[u_0 u_k^{/}] = \psi(1)V\psi(1)^{/} = (I - C)^{-1}V(I - C)^{-1\prime}$ and this matrix can be decomposed into $\Omega = \Delta + \Gamma'$ where $\Delta = \Gamma_0 + \Gamma$, $\Gamma_0 = E[u_0 u_0']$ and $\Gamma = \sum_{k=1}^{\infty} E[u_0 u_k^{/}]$. Let Ω and Δ have elements ω_{ij} and Δ_{ij} (i,j = η,s), respectively. Using these matrices it can be shown that the elements that contribute to the bias and nonnormality of OLS estimates are the quantities $\theta = \omega_{s\eta}/\omega_{ss}$ and $\Delta_{s\eta}$, which measure the long-run correlation and endogeneity between u_{ft} and u_{st}. If these elements are zero then the OLS estimates are asymptotically (mixed) normal, unbiased and efficient[5]. Using the triangular system (13) some tedious calculations show (see Zivot (1995))

$$\theta = (\alpha_s\sigma_{\eta\eta}/(1 - \phi) + \sigma_{\eta\epsilon}/(1 - \phi))\cdot(\alpha_s^2\sigma_{\eta\eta}/(1 - \phi) + 2\alpha_s\sigma_{\eta s}/(1 - \phi) + \sigma_{ss})^{-1},$$

$$\Delta_{s\eta} = \alpha_s\sigma_{\eta\eta}\phi[(1 - \phi)(1 - \phi^2)]^{-1} + \sigma_{\eta s}/(1 - \phi),$$

and these quantities are zero if $\alpha_s = 0$ (spot rates are weakly exogenous for β_s) and $\sigma_{\eta s} = 0$. In typical exchange rate data, however, $\sigma_{\eta s} \approx 0$ and $\sigma_{\eta\eta}$ is much smaller than σ_{ss} which implies that $\theta \approx 0$ and $\Delta_{s\eta} \approx 0$ so the OLS bias is expected to be very small.

To illustrate the expected magnitude of the OLS bias we conducted a simple Monte Carlo experiment where data was generated by (13) with $\beta_s = 1$, $\alpha_s = 1, -3$, $\phi = 0.9$, $\sigma_{ss} = (0.035)^2$, $\sigma_{\eta\eta} = (0.001)^2$, and $\rho_{\eta s} = 0$[6]. When $\alpha_s = 1$ the FRUH is true and when $\alpha_s = -3$ it is not. In both cases the forward premium is highly autocorrelated. Table 2 gives the results of OLS applied to the levels regression (12a) for samples of size T=100 and T=250. In both cases the magnitude of the OLS bias is negligible. The OLS standard errors, however, are biased which cause the size distortions in the nominal 5% *t*-tests of the hypothesis that $\beta_s = 1$.

4.3 Empirical Example

Table 3 presents estimation results for the VAR model (7) and Table 4 gives the results for the triangular model (13) imposing $\beta_s = 1$ for the pound, yen and Canadian dollar monthly exchange rate series. The VAR(1) model was selected for all currencies by likelihood ratio tests for lag lengths and standard model selection criteria. For all currencies, f_t and s_t behave very similarly: the estimated intercepts and error variances are nearly identical and the estimated correlation between ϵ_{ft} and ϵ_{st}, ρ_{fs}, is 0.99. The estimated coefficients from the triangular model essentially mimic the corresponding coefficients from the VAR(1). Table 5 gives the results of Johansen's likelihood ratio test for the number of cointegrating vectors for the three exchange rate series based on the estimation of (7). If the intercepts are restricted to the error correction term then the Johansen rank test finds one cointegrating vector in all cases. If the intercept is unrestricted then the rank tests finds that spot and forward rates for the pound and Canadian dollar are $I(0)$. However, for each series, the likelihood ratio statistic does not reject the hypothesis that the intercepts be restricted to the error correction term. Table 5 also reports the results of the Engle-Granger (1987) two-step residual based ADF *t*-test for no cointegration based on estimating β_s by OLS. For long lag lengths the null of no-cointegration between forward and spot rates is not rejected at the 10% level but for short lag lengths the null is rejected at the 5% level[7]. Table 6 reports estimates of β_s using OLS, Stock and Watson's (1993) dynamic OLS (DOLS) and dynamic GLS (DGLS) lead-lag estimator and Johansen's (1995) reduced rank MLE [8]. The latter three estimators are asymptotically efficient estimators and yield asymptotically valid standard errors. Notice that all of the estimates of β_s are extremely close to 1 and the hypothesis that $\beta_s = 1$ cannot be rejected using the asymptotic t-tests based on DOLS/DGLS and MLE. Table 7 shows the MLEs of the parameters of the VECM (8) where the intercepts are restricted to the error correction term. Notice also that the estimates of α_f and α_s are both significantly negative and of about the same magnitude indicting that the error correction term, which is essentially the forward premium since $\beta_s \approx 1$, is very highly autocorrelated. Further, since the estimates of α_f and α_s are significantly different from zero neither spot nor forward rates appear to be weakly exogenous with respect to the cointegrating parameters.

The above empirical results are based on a two-step procedure of first testing for cointegration between f_t and s_t and then testing if the cointegrating vector is (1,-1). Alternatively, one can use a one-step procedure to test the null of no-cointegration against the joint hypothesis of cointegration with a prespecified cointegrating vector. The advantage of using tests that impose a prespecified cointegrating vector is that if the cointegrating vector is true then the test can have substantially higher power than tests that implicitly estimate the cointegrating vector[9]. The most commonly used one-step procedure to test the joint hypothesis of cointegration between f_t and s_t with the prespecified cointegrating vector (1,-1) is to run either a unit root test (e.g. ADF t-test) or a stationarity test (e.g KPSS test) on the forward premium $f_t - s_t$. Table 5 also reports the ADF unit root tests and KPSS stationarity tests on the forward premia for the three exchange rate series. For short lag lengths the ADF tests indicate that the forward premia are $I(0)$ whereas for long lags the series appear $I(1)$. The KPSS tests give mixed results[10].

Use of the ADF test as a test for no-cointegration against the alternative of cointegration with a prespecified cointegrating vector, however, has been criticized by Kremers, Ericsson and Dolado (1992) and Zivot (1998) as having low power since the ADF test places unrealistic parametric restrictions on the short-run dynamics of the data. They show that a single equation conditional ECM test, based on models similar to (10a) and (10b), can have substantially higher power than the ADF test. A limitation of the conditional ECM test, however, is that it assumes weak exogeneity of one variable with respect to the cointegration parameters. In the context of testing the stationarity of the forward premium it is hard to argue *a priori* that either forward rates or spot rates are weakly exogenous, and the empirical results of table 5 indicate they are not, and so it appears that the conditional ECM test is not appropriate. Horvath and Watson (1995), however, develop a multivariate procedure to test for cointegration with a prespecified cointegrating vector within a VECM that does not require any exogeneity assumptions. The Horvath and Watson test statistic in the present context is simply the Wald statistic for testing the joint hypothesis $\alpha_f = \alpha_s = 0$ where the parameters are estimated by OLS from the VECM (11)[11]. Under the null of no-cointegration the Wald test has a nonstandard asymptotic distribution and Horvath and Watson (1995) supply the appropriate critical values. They show that their test can have considerably higher power than Johansen's rank test, which is based on implicitly estimating the cointegrating vector. Additionally, they show that their test has good power even if the cointegrating vector is moderately misspecified. The estimates of the VECMs imposing the cointegrating vector (1,-1) are presented in Table 8 and the Horvath-Watson Wald statistics are reported in Table 5. Using the Horvath-Watson Wald test, the null of no-cointegration is rejected at the 5% level in favor of the alternative of cointegration with cointegrating vector (1,-1) for all three exchange rates.

Based on the results from Table 8, the FRUH is clearly rejected for all three exchange rate series since the null hypothesis that $\alpha_s = 1$ can be rejected at any reasonable level of significance using an asymptotic *t*-test. As discussed in Engel (1996), a rejection of the FRUH is usually interpreted as evidence for the existence of a time varying risk premium or a peso problem. Using (11b), it is easy to see that the risk premium implied by the VECM (11) is $rp_t^{re} = (1 - \alpha_s)(f_t - s_t)$ and so the forward premium is perfectly correlated with the risk premium.

5. Models of Cointegration between s_{t+1} and f_t implied by cointegration between f_t and s_t

As discussed earlier, cointegration between f_t and s_t with cointegrating vector (1,-1) implies cointegration between s_{t+1} and f_t with cointegrating vector (1,-1). However, the implied VECM and triangular representations for s_{t+1} and f_t based on the simple models of cointegration between f_t and s_t presented in the previous section are somewhat nonstandard. To see this, consider first the derivation of the VECM for Δf_t and Δs_{t+1}. By adding and subtracting $\alpha_s f_{t-1}$ from the right hand side of (11b), led one period, and adding and subtracting $\alpha_f s_t$ from the right hand side of (11a), we get the VECM for $(\Delta f_t, \Delta s_{t+1})'$

$$\Delta f_t = \mu_f - \alpha_f(s_t - f_{t-1}) + \alpha_f \Delta s_t + \epsilon_{ft}, \tag{14a}$$

$$\Delta s_{t+1} = \mu_s - \alpha_s(s_t - f_{t-1}) + \alpha_s \Delta f_t + \epsilon_{s,t+1}. \tag{14b}$$

Notice that in (14) the error correction term is now the lagged forecast error, $s_t - f_{t-1}$, and not the lagged forward premium and that the error terms are separated by one time period and are thus contemporaneously uncorrelated. The errors, however, are not independent due to the correlation between ϵ_{ft} and ϵ_{st}. Consequently, the representation in (14) is not a VECM that can be derived from a finite order VAR model for $(s_{t+1},\ f_t)'$. In addition, since the error correction term enters both equations neither the forward rate nor the future spot rate is weakly exogenous for the cointegration parameters.

Although Δs_t is on the right hand side of (14a) and Δf_t is on the right hand side of (14b) these models should not be interpreted as conditional models since they are derived by simple algebraic manipulation of (11). In particular, since the time index is shifted by one period between the two equations it is more appropriate to interpret Δf_t in (14b) as a predetermined variable. Estimation of (14b) by OLS will yield consistent but not necessarily efficient estimates of α_s [12]. However, estimation of (14a) by OLS is problematic since both $s_t - f_{t-1}$ and Δs_t are correlated with ϵ_{ft}.

Next consider the derivation of the triangular representation. Using (8b) and $f_t - s_t - \mu_c = u_{ft}$ the triangular model for s_{t+1} and f_t becomes

$$s_{t+1} = \mu_c + f_t + \nu_{s,t+1}, \tag{15a}$$

$$\Delta f_t = \nu_{ft}, \tag{15b}$$

where

$$\nu_{s,t+1} = (\alpha_s - 1)u_{ft} + \epsilon_{s,t+1} \tag{15c}$$

$$\nu_{ft} = \alpha_f u_{f,t-1} + \epsilon_{ft}\ . \tag{15d}$$

From (15a,c), we see that the demeaned forecast error, $s_{t+1} - f_t - \mu_c$ is an AR(1) process with additive noise. The serial correlation in the forecast error will disappear if the FRUH is true or if the forward premium is not autocorrelated. Moreover, if the FRUH is not true the large variance of $\epsilon_{s,t+1}$ relative to u_{ft} will make it difficult to detect the serial correlation in the forecast error. Consequently, tests of FRUH based on testing serial correlation in the forecast error or in the residuals from the levels regression (3) are bound to have low power. Although it may be difficult to detect serial correlation in the forecast error, the representation in (15a) shows that the forward premium can be used to help predict the forecast error unless the FRUH is true. Since u_{ft} and $\epsilon_{s,t+1}$ are essentially uncorrelated the AR(1) plus noise process can be given an ARMA(1,1) representation and this implies that the system $(s_{t+1} - f_t\ - \mu_c,\ \Delta f_t\ )'$ cannot be given a simple VAR representation.

The representation in (15) has important implications for testing for cointegration between s_{t+1} and f_t as well as for estimating the cointegrating vector from the levels regression (3). Suppose that s_t and f_t are not cointegrated so that $\mu_c = \alpha_s = 0$, $\phi = 1$ and so $u_{ft} \sim I(1)$. Then (15a) shows that the forecast error can be decomposed into a random walk component, u_{ft}, and an independent stationary component, $\epsilon_{s,t+1}$. Further, the variance of the random walk component is considerably smaller than the variance of the stationary component. In this case, it will be very difficult to detect the random walk component using standard unit root tests on the forecast error. It follows that unit root tests on the forecast error will likely suffer from size distortions and stationarity tests will suffer from low power[13]. Unit root tests on the forward premium, however, do not suffer from such size distortions although they generally will have low power due to the large persistence in the forward premium. To illustrate, Table 9 reports unit root and stationarity tests on the forecast error $s_{t+1} - f_t$ for the three exchange rate series. For short lags the unit root null is strongly rejected and for the long lags the null is only weakly rejected. The null of stationarity is not rejected for all series using the KPSS test.

Now suppose that s_{t+1} and f_t are cointegrated and (15a) is the correct representation given that $\beta_f = 1$. Then the OLS estimate of β_f from the levels regression (3) will be consistent but asymptotically biased and inefficient due to dynamic behavior and feedback between the elements of $v_{t+1} = (v_{s,t+1}, v_{ft})'$. Since the VECM (14) cannot be derived from a finite order cointegrated VAR model for (s_{t+1}, f_t), testing for cointegration and estimating the cointegrating vector using standard VAR techniques is problematic. In particular, since $s_t - f_{t-1}$ is correlated with ϵ_{ft}, testing and estimation methods based on naive VECMs for Δs_{t+1} and Δf_t, like the Horvath-Watson Wald test and the Johansen MLE, are likely to suffer from biases. The dynamic OLS/GLS estimator of Stock and Watson (1993), however, should work well since it is designed to pick-up feedback effects through the inclusion of leads and lags of Δf_t in the levels regression (3). The results of Table 10 show that this is indeed the case for the pound, yen and Canadian dollar. For all series, the OLS estimates of β_f are downward biased but the Stock-Watson DOLS and DGLS estimates are nearly one.

Somewhat surprisingly, the triangular representation (15) shows that the OLS estimate of β_f will be biased even if $\alpha_s = 1$ (the FRUH is true). This result is due to the fact that the long-run covariance matrix of v_{t+1} is not diagonal. To illustrate, let the FRUH be true and suppose that $\alpha_f =$

0 so that the forward premium is not autocorrelated (this assumption greatly simplifies the calculations but does not qualitatively affect the end result). Then the triangular representation (15) simplifies to

$$s_{t+1} = f_t + \epsilon_{s,t+1},$$
$$f_t = f_{t-1} + \epsilon_{ft},$$

and $v_{t+1} = e_{t+1} = (\epsilon_{s,t+1}, \epsilon_{ft})'$. Then by straightforward calculations the long-run covariance matrix of e_{t+1} and its components are

$$\Omega = \begin{bmatrix} \sigma_{ss} & \sigma_{sf} \\ \sigma_{sf} & \sigma_{ff} \end{bmatrix},\ \Gamma_0 = \begin{bmatrix} \sigma_{ss} & 0 \\ 0 & \sigma_{ff} \end{bmatrix},\ \Gamma = \begin{bmatrix} 0 & \sigma_{sf} \\ 0 & 0 \end{bmatrix},\ \Delta = \begin{bmatrix} \sigma_{ss} & \sigma_{sf} \\ 0 & \sigma_{ff} \end{bmatrix},$$

so that $\theta = \sigma_{sf}/\sigma_{ff}$ and $\Delta_{fs} = 0$. Further, since $\sigma_{ss} \approx \sigma_{ff}$ and $\rho_{sf} \approx 1$ it follows that $\theta \approx 1$ and so OLS on the levels regression (3) will suffer from bias even if the FRUH is true. To illustrate the magnitude of the bias, Table 2(c) reports OLS estimates of the levels regression (3) when data are generated from (12). The OLS estimate of β_f is biased downward, to a similar degree observed in empirical results, and the finite sample distribution is heavily left-skewed. The *t*-statistic for testing $\beta_f = 1$ is centered around -1.5 and a nominal 5% test rejects the null that $\beta_f = 1$ about 30% of the time when the null is true. Table 2 also reports results for the Stock-Watson DOLS estimator. In all cases, the Stock-Watson estimator is essentially equal to the true value of unity and the *t*-statistic for testing $\beta_f = 1$ is roughly symmetric and centered around zero. However, there is moderate size distortion in the nominal 5% *t*-tests of $\beta_f = 1$ for T=100 but the distortions dissipates as T increases.

6. A Reinterpretation of Some Recent Results Regarding the FRUH.

The results of the previous sections can be used to reinterpret the results of Norrbin and Reffett (1996), hereafter NR, and Naka and Whitney (1995), hereafter NW, who use particular cointegrated models for $(s_{t+1}, f_t)'$ and find support for the FRUH.

6.1 Norrbin and Reffett's model

NR based their analysis on the following VECM for $(s_{t+1}, f_t)'$

$$\Delta s_{t+1} = \mu_s + \delta_s(s_t - \beta_f f_{t-1}) + \varsigma_{st+1}, \tag{16a}$$

$$\Delta f_t = \mu_f + \delta_f(s_t - \beta_f f_{t-1}) + \varsigma_{ft,} \tag{16b}$$

which is based on a cointegrated VAR(1) model for $(s_{t+1}, f_t)'$[14]. NR are primarily interested in directly testing the LRFRUC, i.e., that (s_{t+1}, f_t) are cointegrated with cointegrating vector (1, -1), and not the FRUH. Their approach is to impose $\beta_f = 1$, estimate (16) by OLS and test the significance of the error correction coefficients δ_s and δ_f. Table 11 presents the estimation results for (16) applied to our data. They find that estimates of δ_s are not statistically different from zero, estimates of δ_f are not statistically different from 1, the R^2s from (16a) and (16b) are close to zero and one, respectively, and the error term from (16b) is highly serially correlated[15]. Our results are very similar. From these results they conclude that s_{t+1} and f_t are cointegrated with cointegrating vector (1,-1) (since $\delta_f \neq 0$) and that spot rates are weakly exogenous for the cointegrating parameters (since $\delta_s = 0$)[16]. Based on their finding that spot rates are weakly exogenous they argue that tests of the LRFRUC constructed from an error correction equation for Δs_{t+1} are bound to lead one to mistakenly reject the LRFRUC and, therefore, reject the FRUH.

Using the cointegrated model for s_{t+1} and f_t implied by the cointegrated model for s_t and f_t presented in section 5 we can give alternative interpretations of NR's results. Most importantly, our results show that NR's claim that spot rates are weakly exogenous is inconsistent with the FRUH. To see how NR arrived at their results observe that (16) is a restricted version of (14) since Δf_t is omitted from (16a) and Δs_t is omitted from (16b). Now, NR's finding that estimates of δ_s are close to zero can be explained by omitted variables bias. For example, if (14) is the true model, with $\mu_f = \mu_s = 0$, then straightforward calculations based on the stylized facts of the data show $plim\ T^{-1}\sum_1^T (s_t - f_{t-1})\Delta f_t \approx \sigma^2$, $plim\ T^{-1}\sum_1^T (s_t - f_{t-1})^2 \approx \sigma^2$ and so $plim\ \hat{\delta}_s \approx 0$. Also, as mentioned in the last section, estimation of (16b) by OLS is problematic due to the correlation between $\delta_f(s_t - \beta_f f_{t-1})$ and e_{ft}. Furthermore, the finding that $\hat{\delta}_f = 1$ with $\beta_f = 1$ in (16b) implies that $\Delta f_t = s_t - f_{t-1} + \hat{\varsigma}_{ft}$ or, equivalently, that $f_t = s_t + \hat{\varsigma}_{ft}$. This combined with the result that $R^2 = 1$ and $\hat{\varsigma}_{ft}$ is highly autocorrelated simply shows that the forward premium is highly autocorrelated and does not provide evidence one way or another about the FRUH. Finally, consider NR's Table 2 which gives the results for the estimation of the error correction model

$$\Delta s_{t+1} = \mu + \alpha(s_t - f_{t-1}) + \delta\Delta f_t + \pi\Delta f_{t-1} + \gamma\Delta s_t + u_{t+1} \tag{17}$$

which mimics the ECM estimated by Hakkio and Rush (1989). NR claim that this regression is misspecified since it mistakenly assumes that forward rates are weakly exogenous (presumably due to the presence of Δf_t). However (17) is in the form of (14b) which is not a conditional model and does not make any assumptions about the weak exogeneity of forward rates so NR's claim is not true.

6.2 Naka and Whitney's model

NW are interested in testing the LRFRUC and the FRUH simultaneously using the following cointegrated triangular representation for $(s_{t+1}, f_t)'$

$$s_{t+1} = \mu + \beta_f f_t + v_{s,t+1}, \tag{18a}$$

$$\Delta f_t = v_{ft}, \tag{18b}$$

where

$$v_{s,t+1} = \rho v_{st} + w_{s,t+1}, \tag{18c}$$

$$v_{ft} = w_{ft}, \tag{18d}$$

and $w_{s,t+1}$, w_{ft} are *i.i.d.* error terms. Notice that (18a) allows for serial correlation in the "levels regression" but the restriction that f_t is strictly exogenous is imposed in (18b). The VECM derived from (18) is

$$\Delta s_{t+1} = (1 - \rho)\mu - (1 - \rho)(s_t - \beta_f f_{t-1}) + \beta_f \Delta f_t + w_{s,t+1}, \tag{19a}$$

$$\Delta f_t = w_{ft}. \tag{19b}$$

In (19a) the speed of adjustment coefficient is directly related to the correlation in the forecast error and the long-run impact of forward rates on future spot rates (the coefficient on f_t) is restricted to be equal to the short-run effect (the coefficient on Δf_t) . In (19), the FRUH imposes the restrictions $\mu = 0$, $\rho = 0$ and $\beta_f = 1$. NW estimate (19a) by nonlinear least squares (NLS) and report estimates of μ and ρ close to zero and estimates of β_f close to one[17]. Table 12 replicates NW's analysis using our data and we find very similar results. Based on these results, NW cannot reject the FRUH.

The triangular representation (18) used by NW is very similar to the triangular model (15) but with some important differences. In particular, since f_t is assumed to be strictly exogenous and $w_{s,t+1}$ and w_{ft} are assumed to be independent NW's model does not allow for feedback between Δs_t and Δf_t or for contemporaneous correlation between ν_{st} and ν_{ft}. These assumptions imply that (19a) is a properly derived conditional model and so efficient estimation of β_f and ρ via nonlinear least squares can be made. These assumptions, however, place unrealistic restrictions on the dynamics of spot and forward rates. For example, suppose (18) is the correct model and that the FRUH is true so $\mu = 0$, $\beta_f = 1$ and $\rho = 0$. Then the derived VECM for $(f_t, s_t)'$ is

$$\Delta f_t = w_{ft}$$

$$\Delta s_t = (f_{t-1} - s_{t-1}) + w_{st}$$

where w_{ft} and w_{st} are independent. This model implies that $f_{t-1} - s_{t-1} = w_{f,t-1} - w_{s,t-1}$, which is a white noise process, and that Δf_t and Δs_t are uncorrelated. Clearly these results are at odds with the observation that the forward premium is highly autocorrelated and that Δf_t and Δs_t are highly contemporaneously correlated. In addition, NW fail to recognize that assuming f_t is strictly exogenous and w_{st} and w_{ft} are independent the results of Phillips and Loretan (1991) and Phillips (1991) show that OLS on (18a) yields asymptotically efficient estimates of β_f and so there is no efficiency gain in estimating the nonlinear error correction model (19a)[18]. Indeed, the results of Tables 10 and 12 show that the OLS and NLS estimates of β_f are almost identical. Finally, since NW's estimates of β_f are essentially unity, their tests for the significance of ρ in (19a) are roughly equivalent to tests for serial correlation in the forecast error $s_{t+1} - f_t$. Given the remarks in section 5 we know that this test of the FRUH is bound to have low power. In sum, by starting with a simple cointegrated model for s_{t+1} and f_t, NW fail to capture some important dynamics between s_t and f_t that provide information about the validity of the FRUH.

7. Conclusion

In this paper we illustrate some potential pitfalls in modeling the cointegrated behavior of spot and forward exchange rates and we are able to give explanations for some puzzling results that commonly occur in exchange rate regressions used to test the FRUH. We find that a simple first order VECM for s_t and f_t captures the important stylized facts of typical monthly exchange rate data and serves as a natural statistical model for explaining exchange rate behavior. We show that the cointegrated model for s_{t+1} and f_t derived from the VECM for s_t and f_t is not a simple finite order VECM and that estimating a first order VECM for s_{t+1} and f_t can lead to mistaken inferences concerning the exogeneity of spot rates and the unbiasedness of forward rates.

References

Baillie, Richard T. (1989): "Econometric tests of rationality and market efficiency," *Econometric Reviews*, 8(2), 151-186.

_____ (1994): "The long memory of the forward premium," *Journal of International Money and Finance*, 11, 208-219.

_____ and Tim Bollerslev (1989): "Common stochastic trends in a system of exchange rates," *Journal of Finance*, 44, 167-181.

Barnhart, Scott W. and Andrew C. Szakmary (1991): "Testing the unbiased forward rate hypothesis: Evidence on unit roots, co-integration, and stochastic coefficients," *Journal of Financial and Quantitative Analysis*, 26, 245-267.

_____, Robert McNown and Myles S. Wallace (1998): "Some answers to puzzles in testing efficiency of the foreign exchange market," unpublished manuscript, Department of Economics, Clemson University, Clemson, SC.

Barkoulas, John and Christopher F. Baum (1995): "A re-examination of the fragility of evidence from cointegration-based tests of foreign exchange market efficiency," unpublished manuscript, Department of Economics, West Virginia University, Morgantown, WV.

Bekaert, Geert. and Robert J. Hodrick (1993): "On biases in the measurement of foreign exchange risk premiums," *Journal of International Money and Finance*, 12, 115-138.

Banerjee, Anindya, Juan Dolado, John W. Galbraith and David F. Hendry (1993): *Co-integration, error correction, and the econometric analysis of non-stationary data*, Advanced texts in econometrics, Oxford university press, Oxford, UK.

Bilson, John F.O. (1981): "The 'speculative efficiency' hypothesis," *Journal of Business*, 54, 435-51.

Campbell, John Y. and Pierre Perron (1991): "Pitfalls and opportunities: What macroeconomists should know about unit roots and cointegration," *NBER Macroeconomics Annual*. MIT Press, Cambridge, MA.

Clarida, Richard H. and Mark P. Taylor (1997): "The term structure of forward exchange premiums and the forecastability of spot exchange rates: correcting the errors," *The Review of Economics and Statistics*, Vol. LXXIX, No. 3, 353-361.

Corbae, Dean, Kian-Guan Lim and Sam Ouliaris (1992): "On cointegration and tests of forward market unbiasedness," *Review of Economics and Statistics*, 74, 728-732.

Cornell, Bradford (1977): "Spot rates, forward rates and exchange market efficiency," *Journal of Financial Economics*, 5, 55-65.

Crowder, William J. (1994): "Foreign exchange market efficiency and common stochastic trends," *Journal of International Money and Finance*, 13, 551-564.

Diebold, Francis X., Javier Gardeazabal and Kamil Yilmaz (1994): "On cointegration and exchange rate dynamics," *Journal of Finance*, 49, 727-735.

Engel, Charles (1996): "The forward discount anomaly and the risk premium: A survey of recent evidence," *Journal of Empirical Finance*, 3, 123-192.

_____ (1998): "Long-run PPP may not hold after all," unpublished manuscript, Department of Economics, University of Washington, Seattle, WA.

Engle, Robert F. and Clive W. Granger (1987): "Cointegration and error correction: representation, estimation and testing," *Econometrica*, 55, 251-76.

Ericsson, Neil R. and J.S. Irons (1994): *Testing exogeneity*, Oxford University Press, Oxford.

Evans, Martin D.D. and Karen Lewis (1993): "Trends in excess returns in currency and bond markets," *European Economic Review*, 37, 1005-1019.

_____ (1995): "Do long-term swings in the dollar affect estimates of the risk premia?" *Review of Financial Studies*, Vol. 8, No. 3, 709-742.

Fama, Eugene (1984): "Forward and spot exchange rates," *Journal of Monetary Economics*, 14, 319-338.

Frenkel, Jacob A. (1980): "Exchange rates, prices and money: lessons from the 1920s," *American Economic Review*, 70, 235-242.

Froot, Kenneth A. and Jeffrey A. Frankel (1989): "Forward discount bias: is it an exchange risk premium?," *Quarterly Journal of Economics*, 104, 139-61.

Godbout, Marie-Josee and Simon van Norden (1996): "Unit root tests and excess returns," Working Paper 96-10, Bank of Canada, Ottawa, ON, Canada.

Granger, Clive W.J. and Paul Newbold (1974): "Spurious regressions in econometrics," *Journal of Econometrics*, 2, 111-20.

Hai, Weike, Nelson Mark and Yangru Wu (1997): "Understanding spot and forward exchange rate regressions," *Journal of Applied Econometrics*, Vol. 12, No. 6, 715-734.

Hakkio, Craig S. and Mark Rush (1989): "Market efficiency and cointegration: An application to the sterling and Deutschmark exchange markets," *Journal of International Money and Finance*, 8, 75-88.

Hamilton, James (1993): *Time series analysis*, Princeton university press, Princeton, NJ.

Horvath, Michael T.K. and Mark W. Watson (1995): "Testing for cointegration when some of the cointegrating vectors are prespecified," *Econometric Theory*, 11, 984-1015.

Isard, Peter (1995): *Exchange rate economics*, Cambridge Surveys of Economic Literature, Cambridge University Press.

Johansen, Soren (1992), "Cointegration in partial systems and the efficiency of single equation analysis," *Journal of Econometrics*, 52, 389-402.

Johansen, Soren (1995): *Likelihood based inference in cointegrated vector autoregressive models*, Oxford University Press, Oxford.

Kremers, Jeroen J.M., Neil R. Ericsson and Juan J. Dolado (1992): "The power of cointegration tests," *Oxford Bulletin of Economics and Statistics*, 54, 325-48.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt and Yongcheol Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," *Journal of Econometrics*, 54, 159-178.

Levich, Richard (1979): "On the efficiency of markets for foreign exchange," in Dornbusch R., Frenkel, J. (Eds.), *International economic policy theory and evidence*. John Hopkins Press, 246 -267,

Liu, Peter C. and G.S. Maddala (1992): "Rationality of survey data and tests for market efficiency in the foreign exchange markets," *Journal of International Money and Finance*, 11, 366-381.

Liuntel, K. B. and K. Paudyal (1998): "Common stochastic trends between forward and spot exchange rates," *Journal of International Money and Finance*, 17, 279-297.

Mark, Nelson C. (1990): "Real and nominal exchange rates in the long run: an empirical investigation," *Journal of International Economics*, 28, 115-136.

Meese, Richard (1989): "Empirical assessment of exchange rate risk premiums," in Stone (ed.) *Financial risk: Theory, evidence and implications*, Kluwer, Boston, MA.

_____, and Kenneth J. Singleton (1982): “On unit roots and the empirical modeling of exchange rates,” *Journal of Finance*, 37, 1029-1035.

Naka, Atsuyuki and Gerald Whitney (1995): “The unbiased forward rate hypothesis re-examined,” *Journal of International Money and Finance*, 14, 857-867.

Newbold, Paul E., Mark E. Wohar, Tony Rayner, Neil Kellard and Christine Ennew (1996), “Two puzzles in the analysis of foreign exchange market efficiency,” unpublished manuscript, Department of Economics, University of Nottingham, Nottingham, UK.

Norrbin, Stefan C. and Kevin L. Reffett (1996): “Exogeneity and forward rate unbiasedness,” *Journal of International Money and Finance*, 15, 267-274.

Park, Joon Y. (1992): “Canonical cointegrating regressions,” *Econometrica*, 60, 119-43.

Phillips, Peter C.B. (1991): “Optimal inference in cointegrated systems,” *Econometrica*, 59, 283-306.

_____ and Bruce Hansen (1990): “Statistical inference in instrumental variables regression with I(1) processes,” *The Review of Economic Studies*, 57, 99-125.

_____ and Mico Loretan (1991): “Estimating long-run economic equilibria,” *Review of Economic Studies*, 57, 99-125.

Siklos, Pierre L. and Clive W. Granger (1996): “Temporary cointegration with an application to interest rate parity,” UCSD Discussion Paper No. 96-11.

Stock, James H. and Mark W. Watson (1993): “A simple estimator of cointegrating vectors in higher order integrated systems,” *Econometrica*, 61, 1097-1107.

Urbain, Jean-Pierre (1993): *Exogeneity in error correction models*, Lecture notes in economics and mathematical systems No. 398, Springer-Verlag, Berlin, Germany.

Watson, M (1995): “VARs and Cointegration,” chapter 47 in Engel, R.F. and D. MacFadden eds. *Handbook of Econometrics*, Vol. IV, North Holland.

Zivot, Eric (1995): “A comment on dynamic specification and testing for unit roots and cointegration,” in K. Hoover (ed.) *Macroeconometrics: Developments, tensions and prospects*, Kluwer, Boston, MA.

_____ (1998): “The power of single equation tests for cointegration when the cointegrating vector is prespecified,” unpublished manuscript, Department of Economics, University of Washington.

Notes

1.Excellent discussions of efficient estimation of cointegrating vectors are given in Phillips and Loretan (1991), Banerjee, Dolado, Galbraith and Hendry (1993), Hamilton (1993), Stock and Watson (1993) and Watson (1995).

2.Given this interpretation of (5), the commonly reported estimates of α_s less than -2 are troubling since it indicates that the single equation error correction model is not stable. This result highlights the need to look at the vector error correction model for $(s_t, f_t)'$.

3.The data are end of month, average of bid and ask rates. All data begin in January 1976, except for forward rates for the Japanese yet which begin in June 1978. All data go through June 1996. The exchange rates obtained are all in terms of British pounds, but were converted to dollar exchange rates.

4.Baillie and Bollerslev (1994), Diebold, Gardeazabal and Yilmaz (1994), Barkoulas and Baum (1995) and Luintel and Paudyal (1998) have stressed the importance of the treatment of the constant term in cointegrated models for spot and forward exchange rates. The restriction on the constant is easily tested with a likelihood ratio test using the Johansen methodology.

5.The Stock-Watson DOLS/DGLS and Johansen ML estimators of β_s asymptotically remove the effects of θ and $\Delta_{s\eta}$ and so are asymptotically unbiased and efficient.

6.The parameters for the Monte Carlo experiment were calibrated from monthly data on UK spot and forward rates quoted in US dollars.

7.This result has been observed by Crowder (1994).

8.The Stock-Watson dynamic OLS and GLS estimators have been used to estimate the cointegrating vector in the levels regression (3) by Evans and Lewis (1993, 1995), Hai, Mark and Wu (1996) and Godbout and van Norden (1996). The Johansen reduced rank estimator has been used by Baillie and Bollerslev (1989), Crowder (1994) and Godbout and van Norden (1996). Other efficient estimators of the cointegrating vector based on nonparametric corrections for long-run correlation and endogeneity include Phillips and Hansen's (1990) FM-OLS estimator and Park's (1992) CCR estimator. Corbae, Lim and Ouliaris (1992) use Park's CCR estimator to investigate the FRUH.

9.See Horvath and Watson (1995) and Zivot (1998).

10. Engel (1996) surveys the empirical evidence on the stationarity of the forward premium and the results are somewhat mixed and depend on the testing procedure, the data frequency and time period. In general, the high persistence and nonhomogeneity of the forward premium reduce the power of unit root tests and distort the size of stationarity tests. In addition, in daily data the forward premium exhibits strong GARCH effects and nonnormality. These problems have led some authors to consider non-standard models of cointegration between f_t and s_t. For example,

Baillie and Bollerslev (1994) consider fractional cointegration, Bekaert and Hodrick (1993), Evans and Lewis (1993, 1995) consider Markov switching cointegration and Siklos and Granger (1996) consider temporary cointegration.

11.Although not considered here, one may also use the Horvath-Watson (1995) test to perform a joint test that the forward premia for all currencies are nonstationary.

12.The estimates will not be efficient because they ignore the restriction that the coefficient on the s_t - f_{t-1} is the same as the coefficient on the Δf_t.

13.A similar point has been made recently by Engel (1998) with regard to testing for a unit root in the real exchange rate.

14.The error term in NR's equations (1b) and (4b) should be ϵ_{i2t-1} not ϵ_{i2t} .

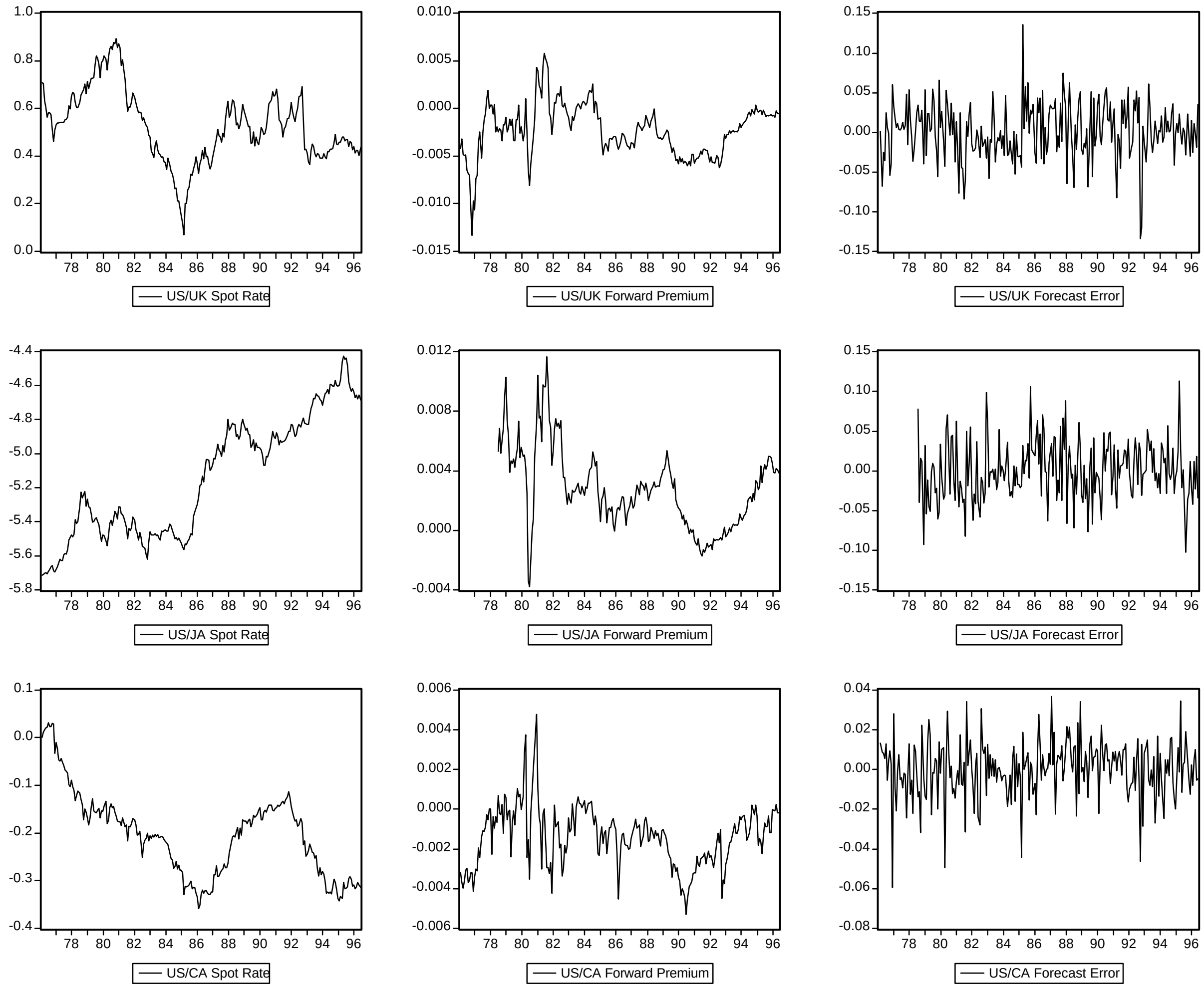
15.Norrbin and Reffett use quarterly data on exchange rates over the period 1973:1 - 1992:4 for the German mark, Canadian dollar, Swiss franc, Japanese Yen and English pound quoted in terms of US dollars.

16.Norrbin and Reffett claim that they do a test of the joint hypothesis that s_{t+1} and f_t are cointegrated with cointegrating vector (1,-1) using the VECM for Δs_{t+1} and Δf_t. The Horvath-Watson Wald test of $\delta_s = \delta_f = 0$ would be the appropriate test statistic. However, since they claim that spot rates are weakly exogenous they base their results on Kremers, Ericsson and Dolado's (1992) single equation conditional error correction model test. But they do not correctly apply the test since they do not estimate a model for Δf_t conditional on Δs_{t+1}.

17.Naka and Whitney (1995) examine monthly exchange rate data covering the period 1974:01 - 1991:04 for the British pound, Canadian dollar, German mark, French franc, Italian lira, Japanese yen and Swiss franc quoted in terms of US dollars.

18.Naka and Whitney (1995) specify that w_{ft} and $w_{s,t+1}$ are *i.i.d.* error terms but they do not make explicit if there is any correlation between w_{ft} and w_{st}. If there is no contemporaneous correlation then OLS is efficient but if these terms are correlated then OLS is not efficient and is asymptotically biased. Moreover, even if w_{ft} and w_{st} are correlated then estimation of Naka and Whitney's nonlinear ECM is not equivalent to maximum likelihood since the long-run covariance matrix in the triangular model is not diagonal.

Figure 1: Monthly Exchange Rate Data



Source: Datastream.

Table 1: Summary Statistics for Exchange Rate Data

	British Pound				Japanese Yen				Canadian Dollar			
	Δs_{t+1}	Δf_{t+1}	f_t - s_t	s_{t+1} - f_t	Δs_{t+1}	Δf_{t+1}	f_t - s_t	s_{t+1} - f_t	Δs_{t+1}	Δf_{t+1}	f_t - s_t	s_{t+1} - f_t
mean	-0.001	-0.001	-0.002	0.001	0.003	0.003	0.003	0.000	-0.001	-0.001	-0.001	0.000
sd	0.034	0.034	0.003	0.035	0.036	0.036	0.003	0.037	0.014	0.014	0.001	0.014
ρ_1	0.087	0.089	0.904	0.111	0.079	0.053	0.926	0.091	-0.108	-0.109	0.786	-0.080
Q	1.859	1.961	201.9***	3.037*	1.537	0.621	186.7***	1.792	2.904*	2.909*	152.6***	1.590
Correlation Matrix	1.000	0.999	-0.135	0.997	1.000	0.999	-0.199	0.997	1.000	0.998	-0.148	0.995
		1.000	-0.143	0.997		1.000	-0.205	0.997		1.000	-0.169	0.995
			1.000	-0.212			1.000	-0.270			1.000	-0.249
				1.000				1.000				1.000

Notes: ρ_1 denotes the first order autocorrelation coefficient and Q denotes the modified Jarque-Berra Q-statistic. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Table 2 : Monte Carlo Estimates of Bias in Levels Regressions

$$f_t = s_t + u_{ft},\ u_{ft} = 0.9u_{ft-1} + \eta_t$$
$$s_t = s_{t-1} + u_{st},\ u_{st} = \alpha_s u_{ft-1} + \epsilon_{st}$$
$$\begin{pmatrix} \eta_t \\ \epsilon_{st} \end{pmatrix} = iid\ N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.001 & 0 \\ 0 & 0.05 \end{pmatrix}\right]$$

(a) Estimated regression: $f_t = a + bs_t + e_t$								
	T = 100				T = 250			
	a	$t_{a=0}$	b	$t_{b=1}$	a	$t_{a=0}$	b	$t_{b=1}$
α_s = 1	-0.000	-0.005 (.642)	1.000	-0.079 (.605)	0.000	0.023 (.653)	1.000	-0.122 (.629)
α_s = -3	-0.000	-0.018 (.644)	1.000	-0.285 (.595)	-0.000	-0.065 (.644)	1.000	-0.303 (.622)
(b) Estimated regression: $s_{t+1} = a + bf_t + e_{t+1}$								
	T = 100				T = 250			
	a	$t_{a=0}$	b	$t_{b=1}$	a	$t_{a=0}$	b	$t_{b=1}$
α_s = 1	-0.000	-0.030 (.275)	0.948	-1.531 (.293)	-0.000	-0.016 (.267)	0.979	-1.518 (.301)
α_s = -3	-0.000	-0.007 (.309)	0.955	-1.415 (.278)	-0.000	-0.014 (.303)	0.983	-1.388 (.265)
(c) Estimated regression: $s_{t+1} = a + bf_t + \sum_{k=-3}^{3} \gamma_k \Delta f_{t-k} + e_{t+1}$								
	T = 100				T = 250			
	a	$t_{a=0}$	b	$t_{b=1}$	a	$t_{a=0}$	b	$t_{b=1}$
α_s = 1	-0.000	-0.001 (.171)	1.000	0.035 (.141)	0.000	0.006 (.108)	1.000	0.017 (.089)
α_s = -3	-0.000	0.005 (.182)	0.999	-0.232 (.142)	-0.000	-0.020 (.115)	1.000	-0.205 (.088)

Notes: Number of simulations = 10,000. Simulations were computed in GAUSS 3.2.14. The values in parentheses indicate the empirical rejection frequency of nominal 5% two-sided tests using asymptotic normal critical values. The standard errors for the Stock-Watson DOLS estimates were computed using an autoregressive estimate of the long-run variance as described in Hamilton (1993) page 610.

Table 3: Bivariate VAR(1) Estimates

$\Delta y_t = \mu + \Pi y_{t-1} + \epsilon_t,\ y_t = (f_t, s_t)',\ \epsilon_t = (\epsilon_{ft}, \epsilon_{st})'$

		Equation	
Currency	Variable/Statistic	Δf_t	Δs_t
British Pound 1976.03 - 1996.06 T = 244	f_{t-1}	-1.771 (0.794)	-1.676 (0.794)
	s_{t-1}	1.744 (0.794)	1.649 (0.795)
	constant	0.009 (0.008)	0.009 (0.008)
	R^2	0.034	0.033
	σ	0.034	0.034
	ρ_{fs}	0.999	
Japanese Yen 1978.08 - 1996.06 T = 215	f_{t-1}	-3.250 (0.944)	-3.178 (0.946)
	s_{t-1}	3.237 (0.941)	3.165 (0.943)
	constant	-0.053 (0.039)	-0.053 (0.039)
	R^2	0.054	0.052
	σ	0.035	0.035
	ρ_{fs}	0.999	
Canadian Dollar 1976.03 - 1996.06 T = 244	f_{t-1}	-2.030 (0.609)	-1.810 (0.608)
	s_{t-1}	2.001 (0.607)	1.782 (0.605)
	constant	-0.010 (0.003)	-0.010 (0.003)
	R^2	0.059	0.051
	σ	0.014	0.014
	ρ_{fs}	0.998	

Notes: Standard errors are in parentheses. ρ_{fs} denotes the correlation between ϵ_f and ϵ_s.

Table 4: Bivariate Triangular Model Estimates With β_s = 1

$u_t = Cu_{t-1} + e_t$, $u_t = (u_{ft}, u_{st})' = (f_t - s_t - \mu_c, \Delta s_t)'$, $e_t = (\eta_t, \epsilon_{st})'$

		Equation	
Currency	Variable/Statistic	u_{ft}	u_{st}
Pound 1976.03 - 1996.06 *T* = 244	u_{ft-1}	0.911 (0.028)	-1.572 (0.808)
	u_{st-1}	0.002 (0.002)	0.070 (0.065)
	R^2	0.822	0.022
	σ	0.001	0.034
	$\rho_{\eta s}$	-0.055	
Yen 1978.08 - 1996.06 *T* = 215	f_{t-1}	0.916 (0.026)	-2.511 (0.911)
	s_{t-1}	-0.004 (0.002)	0.048 (0.068)
	R^2	0.864	0.036
	σ	0.001	0.035
	$\rho_{\eta s}$	-0.095	
CA Dollar 1976.03 - 1996.06 *T* = 244	f_{t-1}	0.794 (0.039)	-1.475 (0.600)
	s_{t-1}	0.008 (0.004)	-0.116 (0.064)
	R^2	0.627	0.025
	σ	0.001	0.014
	$\rho_{\eta s}$	0.033	

Notes: Standard errors are in parentheses. $\rho_{\eta s}$ denotes the correlation between ϵ_η and ϵ_s.

Table 5: Cointegration Tests on f_t and s_t

	Test Statistics						
	Tests based on estimating β				Tests that impose β = (1, -1)′		
Currency	*CADF*	$^{1}_{max}$	$^{2}_{max}$	*LR*	*KPSS*	*ADF*	*HW*
Pound	-3.47** (0) -2.74 (12)	21.98** 5.16**	22.27** 5.44	0.030	0.294 (5)	-3.64*** (0) -2.77* (12)	9.54*
Yen	-3.74** (2) -2.42 (11)	21.08*** 1.33	22.50*** 1.96	0.044	1.319*** (5)	-3.09** (3) -2.23 (11)	18.17***
CA Dollar	-5.58*** (0) -2.70 (10)	39.99*** 4.35**	42.17*** 6.24	0.366	0.380* (5)	-5.36*** (0) -2.91** (5)	12.57**

Notes: *CADF* denotes the Engle-Granger two-step residual-based *ADF* t-statistic; $^{1}_{max}$ and $^{2}_{max}$ denote the Johansen maximum eigenvalue statistic with the intercept unrestricted and restricted, respectively; ADF denotes the augmented Dickey-Fuller t-statistic; *KPSS* denotes the Kwiatkowski, Phillips, Schmidt and Shin (1992) statistic and *HW* denotes the Horvath-Watson Wald statistic. For the maximum eigenvalue statistic, the first row tests the null of no-cointegration versus the alternative of one cointegrating vector and the second row tests the null of one cointegrating vector versus the alternative of two cointegrating vectors. *LR* denotes the likelihood ratio statistic for testing the hypothesis that intercepts are restricted to the error correction term. The Johansen and Horvath Watson tests are based on a VECM with one lag. The number of lags used for the CADF, KPSS and ADF tests are given in parenthesis. ***, ** and * denote rejection at the 1%, 5% and 10% level, respectively.

Table 6: Estimates of the Cointegrating Vector for $(f_t, s_t)'$

OLS: $f_t = \mu_c + \beta_s s_t + u_{ft}$, DOLS/DGLS: $f_t = \mu + \beta_s \cdot s_t + \sum_{k=-3}^{3} \gamma_k \Delta s_{t-k} + \epsilon_t$

MLE: $\Delta y_t = \alpha(\beta' y_{t-1} - \mu_d) + \epsilon_t$

	OLS		Stock-Watson DOLS		Stock-Watson DGLS		Johansen MLE	
Currency	μ_c	β_s	μ_c	β_s	μ_c	β_s	μ_c	β_s
Pound	-0.002 (0.001)	1.000 (0.001)	-0.003 (0.001)	1.000 (0.002)	-0.003 (0.002)	1.002 (0.004)	0.003 (0.002)	1.000 (0.004)
Yen	-0.012 (0.003)	0.997 (0.001)	-0.011 (0.005)	0.997 (0.001)	0.003 (0.013)	1.000 (0.003)	0.008 (0.009)	0.998 (0.002)
CA Dollar	-0.002 (0.001)	0.996 (0.001)	-0.002 (0.001)	0.996 (0.002)	-0.002 (0.001)	0.996 (0.003)	0.003 (0.001)	0.994 (0.003)

Notes: Standard errors are in parentheses. The OLS standard errors are biased. The DOLS standard errors are computed using a Newey-West correction with lag truncation equal to four and the DGLS estimaters are computed via first order Cochrane-Orcutt. The number of leads and lags for the DOLS/DGLS estimator is the same as in Hai, Mark and Wu (1997).

Table 7: ML Estimates of the VECM for $(f_t, s_t)'$

$\Delta s_{t+1} = \alpha_s(f_t - \beta_s s_t - \mu_c) + \epsilon_{st+1}$

Currency	μ_c	α_s	β_s	$\sigma_{ss}^{1/2}$
Pound	-0.003 (0.003)	-1.672 (0.797)	1.000 (0.004)	0.034
Yen	-0.008 (0.009)	-3.220 (0.890)	0.998 (0.002)	0.036
CA Dollar	-0.003 (0.001)	-1.975 (0.606)	0.994 (0.003)	0.014

$\Delta f_{t+1} = \alpha_f (f_t - \beta_s s_t - \mu_c) + \epsilon_{ft+1}$

Currency	μ_c	α_f	β_s	$\sigma_{ff}^{1/2}$
Pound	-0.003 (0.003)	-1.766 (0.797)	1.000 (0.004)	0.034
Yen	-0.008 (0.009)	-3.291 (0.919)	0.998 (0.002)	0.035
CA Dollar	-0.004 (0.001)	-2.187 (0.607)	0.994 (0.003)	0.014

Table 8: Estimates of the VECM for $(f_t, s_t)'$ imposing β = (1, -1)

$$\Delta s_{t+1} = \mu_s + \alpha_s(f_t - s_t) + \epsilon_{st+1}$$

Currency	μ_s	α_s	$\sigma^{1/2}_{ss}$	R^2	*JB*	*LM*	*ARCH*
Pound	-0.005 (0.003)	-1.696 (0.799)	0.034	0.018	43.51 (0.000)	0.611 (0.655)	4.894 (0.001)
Yen	0.010 (0.003)	-2.642 (0.890)	0.035	0.040	8.133 (0.017)	0.121 (0.017)	0.526 (0.716)
CA Dollar	-0.003 (0.001)	-1.386 (0.596)	0.014	0.022	70.58 (0.000)	0.909 (0.459)	0.562 (0.690)

$$\Delta f_{t+1} = \mu_f + \alpha_f (f_t - s_t) + \epsilon_{ft+1}$$

Currency	μ_f	α_f	$\sigma^{1/2}_{ff}$	R^2	*JB*	*LM*	*ARCH*
Pound	-0.005 (0.003)	-1.790 (0.798)	0.034	0.020	45.07 (0.000)	0.624 (0.646)	4.791 (0.001)
Yen	0.010 (0.003)	-2.716 (0.888)	0.035	0.042	8.852 (0.012)	0.089 (0.986)	0.539 (0.707)
CA Dollar	-0.004 (0.001)	-1.598 (0.597)	0.014	0.029	87.76 (0.000)	0.859 (0.489)	0.563 (0.689)

Notes: Standard errors for estimates and *p*-values for test statistics are in parentheses. *JB* denotes the Jacques-Bera statistic, *LM* denotes the LM test for up to 4th order serial correlation and *ARCH* denotes the LM statistic for up to 4th order ARCH effects.

Table 9: Cointegration Tests on s_{t+1} and f_t

	Test Statistics		
	Tests based on estimating β	Tests that impose β = (1, -1)	
Currency	*CADF*	*KPSS*	*ADF*
Pound	-13.73 (0) -3.69* (10)	0.091 (5)	-13.85*** (0) -3.91*** (10)
Yen	-13.42*** (0) -3.30 (10)	0.212 (5)	-13.40*** (0) -3.30** (10)
CA Dollar	-16.81*** (0) -2.72 (12)	0.309 (5)	-16.85*** (0) -2.82* (12)

Notes: See the notes for Table 5.

Table 10: Estimates of the Cointegrating Vector for $(s_{t+1}, f_t)'$

OLS: $s_{t+1} = \mu_c + \beta_f f_t + u_{st}$

DOLS/DGLS: $s_{t+1} = \mu + \beta_f' f_t + \sum_{k=-3}^{3} \gamma_k \Delta f_{t-k} + \epsilon_{t+1}$

	OLS		Stock-Watson DOLS		Stock-Watson DGLS	
Currency	μ_c	β_f	μ_c	β_f	μ_c	β_f
Pound	0.016 (0.008)	0.971 (0.015)	0.003 (0.001)	0.999 (0.002)	0.004 (0.002)	0.997 (0.004)
Yen	-0.004 (0.039)	0.999 (0.008)	0.009 (0.005)	1.002 (0.001)	-0.006 (0.013)	0.999 (0.003)
CA Dollar	-0.003 (0.002)	0.983 (0.010)	0.002 (0.001)	1.004 (0.002)	0.002 (0.001)	1.003 (0.003)

Notes: Seet notes for Table 6.

Table 11: Norrbin and Reffett's Model

$\Delta s_{t+1} = \mu_s + \delta_s(s_t - f_{t-1}) + \varsigma_{st+1}$

Currency	μ_s	δ_s	$\sigma_{ss}^{1/2}$	R^2	JB	LM	$ARCH$
Pound	-0.001 (0.002)	0.094 (0.063)	0.034	0.01	28.41 (0.000)	1.071 (0.371)	2.565 (0.039)
Yen	0.003 (0.002)	0.070 (0.070)	0.036	0.010	3.473 (0.176)	0.168 (0.954)	0.475 (0.754)
CA Dollar	-0.001 (0.001)	-0.100 (0.063)	0.014	0.010	63.41 (0.000)	0.646 (0.630)	0.470 (0.758)

$\Delta f_t = \mu_f + \delta_f (s_t - f_{t-1}) + \varsigma_{ft}$

Currency	μ_f	δ_f	$\sigma_{ff}^{1/2}$	R^2	JB	LM	$ARCH$
Pound	-0.002 (0.000)	0.983 (0.005)	0.003	0.994	32.11 (0.000)	221.4 (0.000)	50.02 (0.000)
Yen	0.003 (0.000)	0.979 (0.005)	0.003	0.995	19.26 (0.000)	163.76 (0.000)	29.97 (0.000)
CA Dollar	-0.001 (0.000)	0.980 (0.007)	0.001	0.989	12.91 (0.002)	92.62 (0.000)	22.11 (0.000)

Notes: See the notes for Table 8.

Table 12: Naka and Whitney's Model

$$\Delta s_{t+1} = \mu(1\text{-}\rho) - (1\text{-}\rho)(s_t - \beta_f f_{t-1}) + \beta_f \Delta f_t + w_{st+1}$$

Currency	μ_s	ρ	β_f	$\sigma^{½}_{ss}$	R^2	*JB*	*LM*	*ARCH*
Pound	0.020 (0.009)	0.131 (0.067)	0.963 (0.017)	0.034	0.004	9.920 (0.008)	0.417 (0.229)	2.838 (0.025)
Yen	-0.008 (0.043)	0.092 (0.069)	0.999 (0.008)	0.036	-0.029	1.978 (0.372)	0.922 (0.452)	0.491 (0.742)
CA Dollar	-0.003 (0.002)	-0.075 (0.065)	0.983 (0.010)	0.014	-0.023	38.67 (0.000)	0.303 (0.876)	0.801 (0.525)

Notes: See the notes for Table 8. The nonlinear least squares estimates are computed using Eviews 3.1. The correlation coefficient ρ is not constrained to lie between -1 and 1.