Cointegration and forward and spot exchange rate regressions

Eric Zivot *

Department of Economics, University of Washington, Box 353330, Seattle, WA 98195-3330, USA

Abstract

We investigate the relationship between cointegration models of the current spot exchange rate, $s_t$, and the current forward rate, $f_t$, and cointegration models of 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 show that simple models of cointegration between $s_t$ and $f_t$ imply complicated models of cointegration between $s_{t+1}$ and $f_t$. Consequently, standard methods are often inappropriate for modeling the cointegrated behavior of $(s_t, f_t)'$ and we show that the use of such methods can lead to erroneous inferences regarding the FRUH. © 2000 Elsevier Science Ltd. All rights reserved.

JEL classification: C32; F31

Keywords: Cointegration; Exchange rates; Forward rate unbiasedness hypothesis; Weak exogeneity

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

* Tel.: +1-206-543-6715; fax: +1-206-685-7477.
E-mail address: ezivot@u.washington.edu (E. Zivot).
Froot and Frankel (1989), concentrated on the regression of the change in the log spot rate, $\Delta s_{t+1}$, 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 et al. (1997), Norrbin and Reffett (1996), Newbold et al. (1998), Clarida and Taylor (1997), Barnhart et al. (1998) and Liuntel 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 $s_t$ 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 $f_t$. For example, 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 lag 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.

Our approach uncovers several interesting results in the analysis of spot and forward exchange rates using cointegration techniques. 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. Using a 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 $f_t$ and 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. Additionally, we show that the empirical results of Naka and Whitney (1995) and Norrbin and Reffett (1996) supporting the FRUH are driven by the specification of the cointegration model for $(s_{t+1}, f_t)$.

Our analysis and results are complementary to those in Hai et al. (1997) and Mark and Wu (1998). Hai et al., assert cointegration between $s_t$ and $f_t$ with cointegrating vector $(1,-1)$ and cointegration between $s_{t+1}$ and $f_t$ and provide direct statistical evidence for these assertions using a variety of methods. However, they do not discuss the implications for modeling cointegration between $s_t$ and $f_t$ versus modeling cointegration between $s_{t+1}$ and $f_t$ for the analysis of exchange rate behavior. Instead, they posit an economic model of exchange rate behavior based on cointegration between $f_t$ and $s_t$ with cointegrating vector $(1,-1)$ and they show that their model is consistent with observed exchange rate behavior. Mark and Wu (1998) simply
assume that $f_t$ and $s_t$ are cointegrated with cointegrating vector $(1,-1)$ and use a simple VECM, much like one used in this paper, to extract an estimate of the excess return series $f_t - E_t[s_{t+1}]$. Based on this estimated series they investigate whether it can be interpreted as a risk premium or something else.

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. Some technical results concerning cointegration used in the paper are summarized in Appendix A.

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 (1989). 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

$$E_t[s_{t+1}] = f_t,$$  \hspace{1cm} (1)

where $E_t[\cdot]$ denotes expectation conditional on information available at time $t$. The FRUH is usually expressed as the levels relationship

$$s_{t+1} = f_t + \zeta_{t+1},$$  \hspace{1cm} (2)

where $\zeta_{t+1}$ is a random variable (rational expectations forecast error) with $E_t[\zeta_{t+1}] = 0$. 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 Eq. (1) itself.

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

$$s_{t+1} = \mu + \beta_f f_t + u_{t+1}$$  \hspace{1cm} (3)

and the null hypothesis that FRUH is true imposes the restrictions $\mu = 0$, $\beta_f = 1$ and $E_t[u_{t+1}] = 0$. Most studies using Eq. (3) found estimates of $\beta_f$ 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 et al. (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 expectations and risk neutrality. Assuming \( s_t \) and \( f_t \) have unit roots, i.e., \( s_t, f_t \sim I(1) \), (see, for example, Meese 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 \( E_t[u_{t+1}]=0 \).

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

\[
s_{t+1} = f_t - r_{t}^{re} + \varepsilon_{t+1},
\]

where \( r_{t}^{re} = f_t - E_t[s_{t+1}] \) represents the rational expectations risk premium. As long as \( r_{t}^{re} \) 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 \( s_{t+1} \) if the risk premium is correlated with variables in the current information set. More generally, in the absence of rational expectations, \( f_t - E_t[s_{t+1}] \) simply represents an excess expected return.

Several authors, e.g. Meese and Singleton (1982) and Meese (1989) and Isard (1995), have stated that since \( s_t \) and \( f_t \) have unit roots the levels regression Eq. (3) is not a valid regression equation because of the spurious regression problem described in Granger and Newbold (1974). However, it is now well known that 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 Eq. (3) is \( I(0) \), then the OLS estimates from Eq. (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 \( \mu \) an \( \beta_f \) will follow non-standard distributions, see Corbae et al. (1992). Fortunately, the biases inherent in OLS estimation of Eq. (3) can be corrected using a variety of methods, see Watson (1995) for a recent review. For example, Hai et al. (1997) use the Stock and Watson (1993) dynamic OLS estimator on the levels regression Eq. (3) and cannot reject \( \beta_f = 1 \) for a number of currencies.

The second regression equation used to test the FRUH is the “differences equation”

\[
\Delta s_{t+1} = \mu^* + \alpha_s(f_t - s_t) + u_{t+1}^*,
\]

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 Eq. (5), surveyed in Engel (1996), overwhelmingly reject the FRUH. In fact, typical estimates of \( \alpha_s \) 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 Eq. (3). Given that \( s_t, f_t \sim I(1) \), for Eq. (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 Eq. (5) implicitly assume that the forward pre-
mium 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, Eq. (5) can be thought of as one equation
in a particular vector error correction model (VECM) for \( (f_t, s_t)' \).\(^1\) Horvath and Wat-
son (1995) and Clarida and Taylor (1997) use VECM-based tests and provide evi-
dence that \( f_t \) and \( s_t \) are cointegrated with cointegrating vector \((1, -1)\).

As noted by Fama (1984), the negative estimates of \( \alpha_s \) found empirically are
consistent with rational expectations and market efficiency and imply certain restric-
tions on the risk premium. Under rational expectations we may rewrite the differ-
ences regression Eq. (5) as

\[
\Delta s_{t+1} = (f_t - s_t) - rp_t^{re} + \xi_{t+1},
\]

so that Eq. (5) is misspecified if risk neutrality fails. Since all variables in Eq. (6)
are \( I(0) \), if the risk premium is correlated with the forward premium then the OLS
estimate of \( \alpha_s \) in the standard differences regression Eq. (5), which omits the risk
premium, will be biased away from the true value of 1. Hence the negative estimates
of \( \alpha_s \) from Eq. (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 \( \alpha_s \) it must be true that \( \text{cov}(E_t[s_{t+1} - s_t, rp_t^{re}] < 0 \) and \( \text{var}(rp_t^{re}) > \text{Var}(E_t[s_{t+1} - s_t]). \)
Models of the foreign exchange risk premium should capture these two inequalities
to be consistent with observed data. The results summarized in Engel (1996), how-
ever, indicate that many economic models of the risk premium are not consistent
with the observed data.

The tests of the FRUH based on Eqs. (3) and (5) involve cointegration either
between \( s_{t+1} \) and \( f_t \) or \( f_t \) and \( s_t \). 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 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 vice-versa. 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. To pre-
view, 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 Eq. (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

\(^1\) Given this interpretation of Eq. (5), the commonly reported estimates of \( \alpha_s \) 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 \( (s_t, f_t)' \).
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

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 US dollar) on the pound, yen and Canadian dollar taken from Datastream. Fig. 1 shows time plots of \(s_{t+1}, f_t - s_t\) (forward premium), and \(s_{t+1} - f_t\) (realized excess profit/loss from speculation) 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 \(\Delta s_{t+1}\) and \(\Delta f_{t+1}\) 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, \(\Delta s_{t+1}, \Delta f_{t+1}\) 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

As we shall see, 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 (1998) and Mark and Wu (1998). Before presenting the empirical results, we begin this section with a review of the properties of such a model. The general bivariate VAR(1) model is \(y_t=\mu + \Phi y_{t-1} + \epsilon_t\), where \(\epsilon_t \sim iid(0, \Sigma)\) and \(\Sigma\) has elements \(\sigma_{ij}(i,j=1,2)\), 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, \(\Pi\) has rank 1 and there exist

---

2 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 yen 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.
Fig. 1. Monthly exchange rate data. Source: Datastream.
Table 1
Summary statistics for exchange rate data

<table>
<thead>
<tr>
<th></th>
<th>British pound</th>
<th>Japanese yen</th>
<th>Canadian dollar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta s_{t+1}$</td>
<td>$\Delta f_{t+1}$</td>
<td>$f_{t} - s_{t}$</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.002</td>
</tr>
<tr>
<td>SD</td>
<td>0.034</td>
<td>0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>$\rho_{1}$</td>
<td>0.087</td>
<td>0.089</td>
<td>0.904</td>
</tr>
<tr>
<td>$Q$</td>
<td>1.859</td>
<td>1.961</td>
<td>201.9***</td>
</tr>
<tr>
<td>Correlation matrix</td>
<td>1.000</td>
<td>0.999</td>
<td>-0.135</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>-0.143</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>-0.212</td>
<td>1.000</td>
</tr>
</tbody>
</table>

* $\rho_{1}$ 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. Additionally, the sample correlations between $\Delta s_{t+1}$ and $\Delta f_{t}$ for the Pound, Yen and Canadian Dollar rates are 0.086, 0.057 and -0.119, respectively.
2×1 vectors $\alpha$ and $\beta$ such that $\Pi=\alpha \beta'$. Using the normalization $\beta=(1,-\beta)_t'$, Eq. (7) becomes the vector error correction model (VECM)

\[
\begin{align*}
\Delta f_t &= \mu_f + \alpha(f_{t-1} - \beta_s f_{t-1}) + \varepsilon_{ft}, \\
\Delta s_t &= \mu_s + \alpha(s_{t-1} - \beta_s f_{t-1}) + \varepsilon_{st}.
\end{align*}
\]

(8a)

(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 Eqs. (8a) and (8b) 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 $\mu_c$. With the intercepts in Eqs. (8a) and (8b) 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 Eq. (7) by $\beta'$ and rearranging gives

\[
f_t - \beta_s s_t - \mu_c = \phi(f_{t-1} - \beta_s f_{t-1} - \mu_c) + \eta_t,
\]

(9)

where $\phi=1+\beta' \alpha_1 + (\alpha_s - \beta_s \alpha_s)$ and $\eta_t=\beta' \varepsilon_t = \varepsilon_{ft} - \beta_s \varepsilon_{st}$. Since Eq. (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. If $\beta_f=1$ then the forward premium is $I(0)$ and follows an AR(1) process and the VECM Eqs. (8a) and (8b) becomes

\[
\begin{align*}
\Delta f_t &= \mu_f + \alpha(f_{t-1} - s_{t-1}) + \varepsilon_{ft}, \\
\Delta s_t &= \mu_s + \alpha(s_{t-1} - s_{t-1}) + \varepsilon_{st}.
\end{align*}
\]

(10a)

(10b)

Notice that Eq. (10b) is simply the standard differences regression Eq. (5) used to test the FRUH. Further, if $\alpha_f$ and $\alpha_s$ are of the same sign and magnitude then the implied value of $\phi$ in Eq. (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 forward premium from Eq. (9) is $\sigma_{ff} = \sigma_{ff}^2 + \sigma_{ss}^2 - 2 \rho_{fs} (\sigma_{ff} \sigma_{ss})^{1/2}$, where $\rho_{fs}$ denotes the correlation between $\varepsilon_{ft}$ and $\varepsilon_{st}$, and will be very small relative to the variances of $\Delta f_t$ and $\Delta s_t$, given the stylized facts that $\sigma_{ff}^2 = \sigma_{ss}^2$ and $\rho_{fs}=1$.

The FRUH places testable restrictions on the VECM Eqs. (8a) and (8b). Necessary conditions for the FRUH to hold are (i) $s_t$ and $f_t$ are cointegrated (ii) $\beta_f=1$ and (iii) $\mu_c=0$. In addition, the FRUH requires that $\alpha_s=1$ in order for the forecast error in Eq. (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, Eqs. (8a) and (8b), led one period, become

\footnote{Baillie and Bollerslev (1989), Diebold et al. (1994), Barkoulas and Baum (1995) and Liungel 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.}
\[
\begin{align*}
\Delta f_{t+1} &= \alpha_f (f_t - s_t) + \varepsilon_{f,t+1}, \\
\Delta s_{t+1} &= (f_t - s_t) + \varepsilon_{s,t+1}.
\end{align*}
\]

Notice that the FRUH requires that the expected change in the 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 \( \alpha_f \) since \( \phi = 1 + (\alpha_f - 1) = \alpha_f \).

The exogeneity status of spot and forward rates with regard to the cointegrating parameters \( \alpha \) and \( \beta \) was the focus of attention in Norrbin and Reffett (1996). Naka and Whitney (1995) also make exogeneity assumptions about forward and spot rates in their analysis. 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 (1999). If \( f_t \) is weakly exogenous with respect to \( (\alpha_s, \beta_s)' \) then \( \alpha_s = 0 \) and efficient estimation of the cointegrating parameters can be made from the single equation conditional error correction model

\[
\begin{align*}
\Delta s_t &= \mu_s + \alpha_s (f_{t-1} - \beta_s s_{t-1}) + \gamma_s \Delta f_t + v_{st},
\end{align*}
\]

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

\[
\begin{align*}
\Delta f_t &= \mu_f + \alpha_f (f_{t-1} - \beta_s s_{t-1}) + \gamma_f \Delta s_t + v_{ft},
\end{align*}
\]

where \( \gamma_f = \sigma_{ff}^{-1} \sigma_{fs} \) and \( v_{ft} \) is uncorrelated with \( \varepsilon_{st} \).

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 the ECM for \( \Delta s_t \) cannot capture the dynamics of typical data. To see this, suppose that forward rates are weakly exogenous so that \( \alpha_f = 0 \). Since \( \sigma_{ss} \approx \sigma_{ff} \approx \sigma_{sf}^2 \) it follows that \( \gamma_s \approx \gamma_f \approx 1 \). If \( \mu_s = 0, \alpha_s = 1 \) and \( \beta_s = 1 \), then the ECM for \( \Delta s_t \) can be rewritten as

\[ s_t = f_{t-1} + \Delta f_t + v_t = f_t + v_t \]

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 for \( (f_t, s_t)' \) is Phillips (1991) triangular representation, which is similar to the triangular representation of a limited information simultaneous equations model. This representation is convenient for studying the asymptotic properties of cointegrating regressions and Baillie (1989) has advocated its use for testing rational expectations restrictions in cointe-
grated 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

\[
\begin{align*}
    f_t &= \mu_c + \beta_s s_t + u_{ft}, \\
    s_t &= s_{t-1} + u_{st},
\end{align*}
\]

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 bivariate moving average representation \( u_t = \psi(L)e_t \), where \( \psi(L) = \sum_{k=0}^{\infty} \psi_k L^k \{ \psi_k \} \) is 1-summable and \( e_t \) is i.i.d. with mean zero and covariance matrix \( V \). Eq. (12a) models the (structural) cointegrating relationship and Eq. (12b) 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 Eqs. (8a) and (8b) with the restricted constant, it is straightforward to show that the derived triangular representation is given by Eqs. (12a) and (12b) with

\[
\begin{align*}
    u_{ft} &= \phi_t u_{ft-1} + \eta_t, \\
    u_{st} &= \alpha_s u_{st-1} + \varepsilon_{st},
\end{align*}
\]

and \( \phi, \alpha_s, \eta_t \) and \( \varepsilon_{st} \) are as previously defined. Eq. (12c) models the disequilibrium error (which equals the forward premium if \( \beta_s = 1 \)) as an AR(1) process and Eq. (12d) allows the lagged error to affect the change in the spot rate.

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 \( \beta_s \) from Eq. (12a) is asymptotically unbiased, efficient and (mixed) normally distributed 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 \( \beta_s \) and \( u_{ft} \) does not Granger cause \( u_{st} \) and vise-versa). Further details are given in Appendix A as well as a small Monte Carlo experiment illustrating biases in various levels regressions.

### 4.3. Empirical example

Table 2 presents estimation results for the VAR model Eq. (7) and Table 3 gives the results for the triangular model Eqs. (12a), (12b), (12c) and (12d) 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 using a maximum of twelve lags. 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 \( e_{ft} \) and \( e_{st} \), \( \rho_{fs} \), is 0.99. The estimated coefficients from the triangular model essentially mimic the corresponding coefficients from the VAR(1). Table 4 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 Eq. (7). If the intercepts are restricted to the error correction
Table 2
Bivariate VAR(1) estimates $\Delta y_t = \mu + \Pi y_{t-1} + \varepsilon_t$, $y_t = (f_t, s_t)'$, $\varepsilon_t = (\varepsilon_f, \varepsilon_s)'$

<table>
<thead>
<tr>
<th>Currency</th>
<th>Variable/Statistic</th>
<th>$\Delta f_t$</th>
<th>$\Delta s_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>British pound 1976.03 – 1996.06 $T = 244$</td>
<td>$f_{t-1}$</td>
<td>$-1.771 (0.794)$</td>
<td>$-1.676 (0.794)$</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$1.744 (0.794)$</td>
<td>$1.649 (0.795)$</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>$0.009 (0.008)$</td>
<td>$0.009 (0.008)$</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.034$</td>
<td>$0.033$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.034$</td>
<td>$0.033$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{fs}$</td>
<td></td>
<td>$0.999$</td>
</tr>
<tr>
<td>Japanese yen 1978.08–1996.06 $T = 215$</td>
<td>$f_{t-1}$</td>
<td>$-3.250 (0.944)$</td>
<td>$-3.178 (0.946)$</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$3.237 (0.946)$</td>
<td>$3.165 (0.943)$</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>$-0.053 (0.039)$</td>
<td>$-0.053 (0.039)$</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.054$</td>
<td>$0.052$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.035$</td>
<td>$0.035$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{fs}$</td>
<td></td>
<td>$0.999$</td>
</tr>
<tr>
<td>Canadian dollar 1976.03–1996.06 $T = 244$</td>
<td>$f_{t-1}$</td>
<td>$-2.030 (0.609)$</td>
<td>$-1.810 (0.608)$</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$2.001 (0.067)$</td>
<td>$1.782 (0.605)$</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>$-0.010 (0.003)$</td>
<td>$-0.010 (0.003)$</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.059$</td>
<td>$0.051$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.014$</td>
<td>$0.014$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{fs}$</td>
<td></td>
<td>$0.998$</td>
</tr>
</tbody>
</table>

* Standard errors are in parentheses. $\rho_{fs}$ denotes the correlation between $\varepsilon_f$ and $\varepsilon_s$.

term then the Johansen trace test finds one cointegrating vector in all cases. If the intercept is unrestricted then the trace 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 4 also reports the results of the Engle and Granger (1987) two-step residual based ADF $t$-test for no cointegration based on estimating $\beta_s$ 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.4 Table 5 reports estimates of $\beta_s$ using OLS, Stock and Watson’s (1993) dynamic OLS (DOLS) and dynamic GLS (DGLS) lead-lag estimator and Johansen’s (1995) reduced rank MLE.5 The latter three estimators are asymptotically efficient

---

4 This result has been observed by Crowder (1994).
5 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) and Evans (1995), Hai et al. (1997) and Godbout and van Norden (1998). The Johansen reduced rank estimator has been used by Baillie and Bollerslev (1989), Crowder (1994) and Godbout and van Norden (1998). Other efficient estimators of the cointegrating vector based on nonparametric corrections for long-run correlation and endogeneity include Phillips and Hansen (1990) FM–OLS estimator and Park (1992) CCR estimator. Corbae et al. (1992) use Park’s CCR estimator to investigate the FRUH.
Bivariate triangular model estimates with $\beta_z = 1$, $u_t = Cu_{t-1} + \epsilon_t$, $u_t = (u_{f_t}, u_{s_t})' = (f_t - s_t - \mu_s, \Delta s_t)'$, $\epsilon_t = (\eta_t, \epsilon_s)'$.

<table>
<thead>
<tr>
<th>Currency</th>
<th>Variable/Statistic</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u_{f_{t-1}}$</td>
<td>$0.911$ (0.028)</td>
</tr>
<tr>
<td></td>
<td>$u_{s_{t-1}}$</td>
<td>$0.002$ (0.002)</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.822$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.001$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{\eta_f}$</td>
<td>$-0.055$</td>
</tr>
<tr>
<td>Pound 1976.03–1996.06 $T=244$</td>
<td>$f_{t-1}$</td>
<td>$0.916$ (0.026)</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$-0.004$ (0.002)</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.864$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.001$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{\eta_s}$</td>
<td>$-0.095$</td>
</tr>
<tr>
<td>Yen 1978.08–1996.06 $T=215$</td>
<td>$f_{t-1}$</td>
<td>$0.794$ (0.039)</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$0.008$ (0.004)</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.627$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.001$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{\eta_s}$</td>
<td>$0.033$</td>
</tr>
<tr>
<td>CA dollar 1976.03–1996.06 $T=244$</td>
<td>$f_{t-1}$</td>
<td>$0.794$ (0.039)</td>
</tr>
<tr>
<td></td>
<td>$s_{t-1}$</td>
<td>$0.008$ (0.004)</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>$0.627$</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$0.001$</td>
</tr>
<tr>
<td></td>
<td>$\rho_{\eta_s}$</td>
<td>$0.033$</td>
</tr>
</tbody>
</table>

* Standard errors are in parentheses. $\rho_{\eta_s}$ denotes the correlation between $\epsilon_\eta$ and $\epsilon_s$.

Cointegration tests on $f_t$ and $s_t$*

<table>
<thead>
<tr>
<th>Test statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests based on estimating $\beta$</td>
</tr>
<tr>
<td>Currency</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Pound</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Yen</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CA Dollar</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* CADF denotes the Engle–Granger two-step residual-based ADF t-statistic; $\lambda^1_{\text{max}}$ and $\lambda^2_{\text{max}}$ denote the Johansen trace statistic with the intercept unrestricted and restricted, respectively; ADF denotes the augmented Dickey–Fuller t-statistic; KPSS denotes the Kwiatkowski et al. (1992) statistic and HW denotes the Horvath–Watson Wald statistic. For the trace 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 5
Estimates of the cointegrating vector for \((f_t, s_t)\)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Stock–Watson DOLS</th>
<th>Stock–Watson DGLS</th>
<th>Johansen MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>-0.002</td>
<td>1.000</td>
<td>-0.003</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Yen</td>
<td>-0.012</td>
<td>0.997</td>
<td>-0.011</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.005)</td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>CA Dollar</td>
<td>-0.002</td>
<td>0.996</td>
<td>-0.002</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

* 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 estimators are computed via first order Cochrane–Orcutt. The number of leads and lags for the DOLS/DGLS estimator is the same as in Hai et al. (1997).

estimators and yield asymptotically valid standard errors. Notice that all of the estimates of \(\beta_s\) 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 6 shows the MLEs of the parameters of the VECM Eqs. (8a) and (8b) where the intercepts are restricted to the error correction term. Notice that the estimates of \(\alpha_f^\prime\) and \(\alpha_s\) are both significantly negative and of about the same magnitude indicating that the error correction term, which is essentially the forward premium since \(\beta_s = 1\), is very highly autocorrelated. Further, since the estimates of \(\alpha_f^\prime\) and \(\alpha_s\) 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 \((1, -1)\).

Table 6
ML estimates of the VECM for \((f_t, s_t)\)

<table>
<thead>
<tr>
<th>(\Delta s_{t+1} = \alpha_s (f_t - \beta_s s_t - \mu_s) + \varepsilon_{s_{t+1}})</th>
<th>(\alpha_s)</th>
<th>(\beta_s)</th>
<th>(\sigma_{ss}^{1/2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currency</td>
<td>(\mu_s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pound</td>
<td>-0.003(0.003)</td>
<td>-1.672(0.797)</td>
<td>1.000(0.004)</td>
</tr>
<tr>
<td>Yen</td>
<td>-0.008(0.009)</td>
<td>-3.220(0.890)</td>
<td>0.998(0.002)</td>
</tr>
<tr>
<td>CA dollar</td>
<td>-0.003(0.001)</td>
<td>-1.975(0.606)</td>
<td>0.994(0.003)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\Delta f_{t+1} = \alpha_f (f_t - \beta_f s_t - \mu_f) + \varepsilon_{f_{t+1}})</th>
<th>(\mu_f)</th>
<th>(\alpha_f)</th>
<th>(\beta_f)</th>
<th>(\sigma_{ff}^{1/2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currency</td>
<td>(\mu_f)</td>
<td>(\alpha_f)</td>
<td>(\beta_f)</td>
<td>(\sigma_{ff}^{1/2})</td>
</tr>
<tr>
<td>Pound</td>
<td>-0.003(0.003)</td>
<td>-1.776(0.797)</td>
<td>1.000(0.004)</td>
<td>0.034</td>
</tr>
<tr>
<td>Yen</td>
<td>-0.008(0.009)</td>
<td>-3.291(0.919)</td>
<td>0.998(0.002)</td>
<td>0.035</td>
</tr>
<tr>
<td>CA dollar</td>
<td>-0.004(0.001)</td>
<td>-2.187(0.607)</td>
<td>0.994(0.003)</td>
<td>0.014</td>
</tr>
</tbody>
</table>
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. Table 4 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.

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 et al. (1992), Horvath and Watson (1995) and Zivot (1999) as having low power since the ADF test places unrealistic parametric restrictions on the short-run dynamics. Horvath and Watson develop a multivariate procedure to test for cointegration with a prespecified cointegrating vector within a VECM that is particularly well suited to the present context. Their test statistic 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 Eqs. (10a) and (10b). The estimates of the VECMs imposing the cointegrating vector $(1, -1)$ are presented in Table 7 and the Horvath–Watson Wald statistics are reported in Table 4 where we see that 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 7, 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), rejection of the FRUH can be interpreted as evidence for the existence of a time varying risk premium under rational expectations. Now, the excess expected return implied by the VECM Eqs. (10a) and (10b) is $f_r - E[f_{s+1}] = (1 - \alpha_s)(f_r - s_r)$. To determine if this estimate corresponds to a risk premium is beyond the scope of the present

---


7 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 $f_r$ and $s_r$. For example, Baillie (1994) consider fractional cointegration, Bekaert and Hodrick (1993), Evans and Lewis (1993) and Evans (1995) consider Markov switching cointegration and Siklos and Granger (1996) consider temporary cointegration.

8 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. Although not considered here, one may also use the Horvath and Watson (1995) test to perform a joint test of the hypothesis that the forward premia for all currencies are nonstationary.
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 \( \Delta f_t \) and \( \Delta s_{t+1} \). By adding and subtracting \( \alpha_s f_{t-1} \) from the right hand side of Eq. (10b), led one period, and adding and subtracting \( \alpha_f s_t \) from the right hand side of Eq. (10a), we get the VECM for \((\Delta f_t, \Delta s_{t+1})'\)

\[
\Delta f_t = \mu_f - \alpha_f (s_{t-1} - f_{t-1}) + \alpha_s \Delta s_t + \epsilon_f,
\]

\[
\Delta s_{t+1} = \mu_s - \alpha_s (s_{t-1} - f_{t-1}) + \alpha_f \Delta f_t + \epsilon_{s,t+1}.
\]

Notice that in Eqs. (13a) and (13b) the error correction term is now the lagged forecast error, \( s_{t-1} - f_{t-1} \), 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 $e_{ft}$ and $e_{st}$. The representation in Eqs. (13a) and (13b) 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.

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

$$s_{t+1} = \mu_c + f_t + v_{s,t+1}, \quad (14a)$$
$$\Delta f_t = v_{ft}, \quad (14b)$$

where

$$v_{s,t+1} = (\alpha_s - 1)u_{ft} + e_{s,t+1} \quad (14c)$$
$$v_{ft} = \alpha_f u_{ft-1} + e_{ft}. \quad (14d)$$

Recall from Eq. (9) that the demeaned forward premium, $f_t-s_t-\mu_c=u_{ft}$, follows an AR(1) process. Hence, from Eq. (14c), 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 $e_{s,t+1}$ 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 Eq. (3) are bound to have low power. Although it may be difficult to detect serial correlation in the forecast error, the representation in Eq. (14a) shows that the forward premium can be used to help predict the forecast error unless the FRUH is true. Since $u_{ft}$ and $e_{s,t+1}$ are 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 Eqs. (14a), (14b), (14c) and (14d) has important implications for testing cointegration between $s_{t+1}$ and $f_t$. For example, suppose that $s_t$ and $f_t$ are not cointegrated so that $\mu_c=\alpha_s=0$, $\phi=1$ and $u_{ft} \sim 1(1)$. Then Eq. (14a) shows that the forecast error can be decomposed into a random walk component, $u_{ft}$, and an independent stationary component, $e_{s,t+1}$, where 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.\(^9\) Unit root tests on the forward premium, however, do not suffer from

\(^9\) A similar point has been made recently by Engel (1999) with regard to testing for a unit root in the real exchange rate.
such size distortions although they generally will have low power due to the large persistence in the forward premium. To illustrate, Table 8 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.

The representation in Eqs. (14a), (14b), (14c) and (14d) can be used to guide inference based on the levels regression Eq. (3). Suppose that $s_{t+1}$ and $f_t$ are cointegrated and Eq. (14a) is the correct representation given that $\beta_f = 1$. Then the OLS estimate of $\beta_f$ from the levels regression Eq. (3) will be consistent but asymptotically biased, inefficient and non-normal due to dynamic behavior and feedback between the elements of $v_{t+1} = (v_{s,t+1}, v_{f,t})'$. Since the VECM Eqs. (14a), (14b), (14c) and (14d) cannot be derived from a finite order cointegrated VAR model for $(s_{t+1}, f_t)'$, estimating the cointegrating vector using VAR techniques (e.g. Johansen’s mle) is problematic. 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 $\Delta f_t$ in the levels regression Eq. (3). The results of Table 9 show that this is indeed the case for the pound, yen and Canadian dollar. For all series, the OLS estimates of $\beta_f$ are downward biased but the Stock–Watson DOLS and DGLS estimates are nearly one. The downward bias in the OLS estimate of $\beta_f$ can be explained using the triangular model Eqs. (14a), (14b), (14c) and (14d), the details of which are presented in Appendix A.

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.

Table 8
Cointegration tests on $s_{t+1}$ and $f_t$:

<table>
<thead>
<tr>
<th>Currency</th>
<th>Test statistics</th>
<th>Tests based on estimating $\beta$</th>
<th>Tests that impose $\beta = (1, -1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CADF</td>
<td>KPSS</td>
</tr>
<tr>
<td>Pound</td>
<td>$-13.73$ (0)</td>
<td>0.091 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-3.69^*$ (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>$-13.42***$ (0)</td>
<td>0.212 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-3.30$ (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA dollar</td>
<td>$-16.81***$ (0)</td>
<td>0.309 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$-2.72$ (12)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* See the notes for Table 4.
Table 9
Estimates of the cointegrating vector for \((s_{t+1}, f_t)\)^T

<table>
<thead>
<tr>
<th>Currency</th>
<th>OLS</th>
<th>Stock–Watson DOLS</th>
<th>Stock–Watson DGLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\mu_c)</td>
<td>(\beta_f)</td>
<td>(\mu_c)</td>
</tr>
<tr>
<td>Pound</td>
<td>0.016</td>
<td>0.971</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.015)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Yen</td>
<td>-0.004</td>
<td>0.999</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.008)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>CA Dollar</td>
<td>-0.003</td>
<td>0.983</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.010)</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

See notes for Table 5.

6.1. Norrbin and Reffett’s model

NR based their analysis on the following VECM for \((s_{t+1}, f_t)\)^T:

\[
\begin{align*}
\Delta s_{t+1} &= \mu_s + \delta_s (s_t - \beta_f f_{t-1}) + s_{st+1}, \\
\Delta f_t &= \mu_f + \delta_f (s_t - \beta_f f_{t-1}) + s_{ft}.
\end{align*}
\]

(15a) (15b)

which is based on a cointegrated VAR(1) model for \((s_{t+1}, f_t)\)^T. NR are primarily interested in directly testing that \((s_{t+1}, f_t)\)^T are cointegrated with cointegrating vector \((1, -1)\) and not the FRUH. Their approach is to impose \(\beta_f = 1\), estimate Eqs. (15a) and (15b) by OLS and test the significance of the error correction coefficients \(\delta_s\) and \(\delta_f\). Table 10 presents the estimation results for Eqs. (15a) and (15b) applied to our data. They find that estimates of \(\delta_s\) are not statistically different from zero, estimates of \(\delta_f\) are not statistically different from 1, the \(R^2\)s from Eqs. (15a) and (15b) are close to zero and one, respectively, and the error term from Eq. (15b) is highly serially correlated.\(^{11}\) Our results are very similar. NR 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 \(\delta_s = 0\)).\(^{12}\) Based on their finding that spot rates are weakly exogenous they argue that tests constructed from

\(^{10}\) The error term in NR’s equations (1b) and (4b) should be \(\varepsilon_{st+1} \neq \varepsilon_{st}\).

\(^{11}\) 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.

\(^{12}\) 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 \(\Delta s_{t+1}\) and \(\Delta f_t\). 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 et al. (1992) single equation conditional error correction model test. But they do not correctly apply the test since they do not estimate a model for \(\Delta f_t\) conditional on \(\Delta s_{t+1}\).
Table 10
Norrbin and Reffett’s model a

<table>
<thead>
<tr>
<th>Currency</th>
<th>$\Delta s_{t+1} = \mu_s + \delta_s(s_t - f_{t-1}) + \mu_{s_{t+1}}$</th>
<th>$\mu_s$</th>
<th>$\delta_s$</th>
<th>$\sigma_{12}^s$</th>
<th>$R^2$</th>
<th>JB</th>
<th>LM</th>
<th>ARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pound</td>
<td>-0.001</td>
<td>0.094</td>
<td>0.034</td>
<td>0.01</td>
<td>28.41</td>
<td>1.071</td>
<td>2.565</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.063)</td>
<td>(0.000)</td>
<td>(0.371)</td>
<td>(0.039)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.003</td>
<td>0.070</td>
<td>0.036</td>
<td>0.010</td>
<td>3.473</td>
<td>0.168</td>
<td>0.475</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.070)</td>
<td>(0.176)</td>
<td>(0.954)</td>
<td>(0.754)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA dollar</td>
<td>-0.001</td>
<td>-0.100</td>
<td>0.014</td>
<td>0.010</td>
<td>63.41</td>
<td>0.646</td>
<td>0.470</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.063)</td>
<td>(0.000)</td>
<td>(0.630)</td>
<td>(0.758)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Currency</th>
<th>$\Delta f_t = \mu_f + \delta_f(s_t - f_{t-1}) + \sigma_{12}^f$</th>
<th>$\mu_f$</th>
<th>$\delta_f$</th>
<th>$\sigma_{12}^f$</th>
<th>$R^2$</th>
<th>JB</th>
<th>LM</th>
<th>ARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pound</td>
<td>-0.002</td>
<td>0.983</td>
<td>0.003</td>
<td>0.994</td>
<td>32.11</td>
<td>221.4</td>
<td>50.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.005)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>0.003</td>
<td>0.979</td>
<td>0.003</td>
<td>0.995</td>
<td>19.26</td>
<td>163.76</td>
<td>29.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.005)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA dollar</td>
<td>-0.001</td>
<td>0.980</td>
<td>0.001</td>
<td>0.989</td>
<td>12.91</td>
<td>92.62</td>
<td>22.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.007)</td>
<td>(0.002)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a See the notes for Table 7.

an error correction equation for $\Delta s_{t+1}$ (like Eq. (16)) are bound to lead one to mistakenly reject $\beta_f=1$ 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 Eqs. (15a) and (15b) is a restricted version of Eqs. (13a) and (13b) since $\Delta f_t$ is omitted from Eq. (15a) and $\Delta s_t$ is omitted from Eq. (15b). Now, NR’s finding that estimates of $\delta_f$ are close to zero can be explained by omitted variables bias. For example, if Eqs. (13a) and (13b) is the true model, with $m_f = m_s = 0$, then straightforward calculations based on the stylized facts of the data show $T^{-1}\Sigma(s_t - f_{t-1})^2 \approx \sigma^2$, $T^{-1}\Sigma(s_t - f_{t-1})^4 \approx \sigma^4$ and so $\delta_f \approx 0$. Also, as mentioned in the last section, estimation of Eq. (15b) by OLS is problematic due to the correlation between $\delta_f(s_t - \beta_f f_{t-1})$ and $e_{f_t}$. Furthermore, the finding that $\delta_f = 1$ with $\beta_f = 1$ in Eq. (15b) implies that $\Delta f_t = s_t - f_{t-1} + \sigma_n$ or, equivalently, that $f_t = s_t + \sigma_n$. This combined with the result that $R^2 = 1$ and $\sigma_n$ 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

$$\Delta s_{t+1} = \mu + \alpha(s_t - f_{t-1}) + \delta f_t + \pi \Delta f_{t-1} + \gamma \Delta s_t + u_{t+1}$$ (16)

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 $\Delta f_t$). However Eq. (16) is in the form of Eq. (13b) 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 FRUH using the following cointegrated triangular representation for $(s_{t+1}, f_t)’$

\[
s_{t+1} = \mu + \beta_f f_t + \nu_{s,t+1},
\]

\[
\Delta f_t = \nu_{ft},
\]

where

\[
\nu_{s,t+1} = \rho \nu_{st} + \nu_{s,t+1},
\]

\[
\nu_{ft} = w_{ft},
\]

and $\nu_{s,t+1}, w_{ft}$ are i.i.d. error terms. Notice that Eq. (17a) allows for serial correlation in the “levels regression” but the restriction that $f_t$ is strictly exogenous is imposed in Eq. (17b). The VECM derived from Eqs. (17a), (17b), (17c) and (17d) is

\[
\Delta s_{t+1} = (1-\rho)\mu - (1-\rho)(s_t - \beta_f f_{t-1}) + \beta_f \Delta f_t + \nu_{s,t+1},
\]

\[
\Delta f_t = w_{ft}.
\]

In Eq. (18a) the speed of adjustment coefficient on the lagged error correction term 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 $\Delta f_t$). In Eqs. (18a) and (18b), the FRUH imposes the restrictions $\mu=0, \rho=0$ and $\beta_f=1$. NW estimate Eq. (18a) by nonlinear least squares (NLS) and report estimates of $\mu$ and $\rho$ close to zero and estimates of $\beta_f$ close to one.\textsuperscript{13} Table 11 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 Eqs. (17a), (17b), (17c) and (17d) used by NW is very similar to the triangular model Eqs. (14a), (14b), (14c) and (14d) but with some important differences. In particular, since $f_t$ is assumed to be strictly exogenous and $\nu_{s,t+1}$ and $w_{ft}$ are assumed to be independent NW’s model does not allow for feedback between $\Delta s_{t+1}$ and $\Delta f_t$ or for contemporaneous correlation between $\nu_{st}$ and $\nu_{ft}$. These assumptions, however, place unrealistic restrictions on the dynamics of spot and forward rates. For example, suppose Eqs. (17a), (17b), (17c) and (17d) is the correct

Table 11
Naka and Whitney’s model $\Delta s_{t+1} = \mu (1-\rho) - (1-\rho)(s_t - \beta f_{t-1}) + \beta_f \Delta f_t + w_{s,t+1} \text{.}^a$

<table>
<thead>
<tr>
<th>Currency</th>
<th>$\mu_s$</th>
<th>$\rho$</th>
<th>$\beta_f$</th>
<th>$\sigma_{12}^2$</th>
<th>$R^2$</th>
<th>JB</th>
<th>LM</th>
<th>ARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pound</td>
<td>0.020</td>
<td>0.131</td>
<td>0.963</td>
<td>0.034</td>
<td>0.004</td>
<td>9.920</td>
<td>0.417</td>
<td>2.838</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.067)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yen</td>
<td>-0.008</td>
<td>0.092</td>
<td>0.999</td>
<td>0.036</td>
<td>-0.029</td>
<td>1.978</td>
<td>0.922</td>
<td>0.491</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.069)</td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA dollar</td>
<td>-0.003</td>
<td>-0.075</td>
<td>0.983</td>
<td>0.014</td>
<td>-0.023</td>
<td>38.67</td>
<td>0.303</td>
<td>0.801</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.065)</td>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ See the notes for Table 7. The nonlinear least squares estimates are computed using Eviews 3.1.

model and that NW’s claim that the FRUH is true so $\mu=0$, $\beta_f=1$ and $\rho=0$. Then the derived VECM for $(f_t, s_t)$ is

$$
\Delta f_t = w_{f,t},
$$

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

where $w_{f,t}$ and $w_{s,t}$ 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 $\Delta f_t$ and $\Delta s_t$ are uncorrelated. Clearly these results are at odds with the observation that the forward premium is highly autocorrelated and that $\Delta f_t$ and $\Delta s_t$ are highly contemporaneously correlated. In addition, NW fail to recognize that assuming $f_t$ is strictly exogenous and $w_{s,t}$ and $w_{f,t}$ are independent the results of Phillips and Loretan (1991) and Phillips (1991) show that OLS on Eq. (17a) yields asymptotically efficient estimates of $\beta_f$ and so there is no efficiency gain in estimating the nonlinear error correction model Eq. (18a). Indeed, the results of Tables 9 and 11 show that the OLS and NLS estimates of $\beta_f$ and the estimates of the standard errors are almost identical (see also NW’s Tables 1A and 2). Finally, since NW’s estimates of $\beta_f$ are essentially unity, their estimates of $\rho$ in Eq. (18a) are essentially estimates of the first order autocorrelation in the forecast error $s_{t+1} - f_t$ (compare Table 1 with Table 11) and tests for the significance of $\rho$ in are roughly equivalent to tests for serial correlation in the forecast error $s_{t+1} - f_t$. Given the remarks in Section 5 about the decomposition of the forecast error into a small variance AR(1) component and a large variance noise component we know that the test of $\rho=0$ is likely 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.

---

$^{14}$ Naka and Whitney (1995) specify that $w_{f,t}$ and $w_{s,t+1}$ are $i.i.d.$ error terms but they do not make explicit if there is any correlation between $w_{f,t}$ and $w_{s,t}$. 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 $w_{f,t}$ and $w_{s,t}$ 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.
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.

The form of the cointegration model for $(f_t, s_t)'$ determines the estimate of the excess expected return $f_t - E[s_{t+1}]$. Several researchers (e.g. Barnhart and Szakmary, 1991; Naka and Whitney, 1995; Engel, 1996) have noted that the estimates of $\alpha_s$ can vary considerably over sub-samples indicating potential parameter instability in the VECM for $(f_t, s_t)'$. Ignoring this parameter instability, if it exists, can have large influences on the derived estimate of $f_t - E[s_{t+1}]$. The investigation of the causes of parameter instability in VECMs for $(f_t, s_t)'$ and the implications of such instability for derived estimates of the foreign exchange risk premium is the focus of our future research.

Acknowledgements

Thanks to Charles Engel for his insights on the issues presented herein, to two anonymous referees for helpful comments that greatly improved the presentation of results in the paper, and to Vinay Datar for motivating me to write the paper.

Appendix A

This appendix summarizes some technical details concerning biases in simple cointegrating regressions. Consider the triangular model Eqs. (12a), (12b), (12c) and (12d) and define $e_t=(\eta_t, e_{st})'$. Then the vector $u_t=(u_{ft}, u_{st})'=(f_t - \beta_s s_t - \mu, \Delta s_t)'$ has the VAR(1) representation $u_t=Cu_{t-1}+e_t$, where

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

and so $\psi(L)=(I-CL)^{-1}$. As noted previously, $\sigma_{\eta \eta}$ is very small relative to $\sigma_{ss}$ and $\sigma_{\eta \eta}=\sigma_{ss} - \sigma_{ss}$. The correlation and feedback effects that cause OLS estimates in Eq. (12a), for example, to be biased and non-normal, 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 \( \Omega = \sum_{k=-\infty}^{\infty} E[u_{0,t} u_k'] = \psi(1)V\psi(1)' = (I-C)^{-1}V(I-C)^{-1}' \) and this matrix can be decomposed into \( \Omega = \Delta' + \Gamma' \) where \( \Delta = \Gamma_0 + \Gamma, \Gamma_0 = E[u_{0,t} u_t'] \) and \( \Gamma = \sum_{k=1}^{\infty} E[u_{0,t} u_k'] \).

Let \( \Omega \) and \( \Delta \) have elements \( \omega_{ij} \) and \( \Delta_{ij} (i,j = \eta,s) \), respectively. Using these matrices it can be shown that the elements that contribute to the bias and nonnormality of OLS estimates are the quantities \( \theta = \omega_{\eta\eta}/\omega_{ss} \) and \( \Delta_{\eta\eta} \), which measure the long-run correlation and endogeneity between \( u_t \) and \( u_{st} \). If these elements are zero then the OLS estimates are asymptotically unbiased, (mixed) normal and efficient.\(^{15}\)

To illustrate the asymptotic bias resulting from OLS estimation of \( \beta_s \) in Eq. (12a), some tedious calculations show (see Zivot, 1995)

\[
\theta = (\alpha_s \sigma_{\eta\eta}/(1 - \phi) + \sigma_{\eta\eta}/(1 - \phi))^2 - (2 \alpha_s \sigma_{\eta\eta}/(1 - \phi) + \sigma_{ss}/(1 - \phi) + \sigma_{\eta\eta}/(1 - \phi))^{-1},
\]

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

and these quantities are zero if \( \alpha_s = 0 \) (spot rates are weakly exogenous for \( \beta_s \)) and \( \sigma_{\eta\eta} = 0 \). In typical exchange rate data, however, \( \sigma_{\eta\eta} = 0 \) and \( \sigma_{\eta\eta} = 0 \) is much smaller than \( \sigma_{ss} \), which implies that \( \theta \approx 0 \) and \( \Delta_{\eta\eta} \approx 0 \) so the OLS bias for \( \beta_s \) is expected to be very small. To illustrate, we conducted a simple Monte Carlo experiment where data was generated by Eqs. (12a), (12b), (12c) and (12d) with \( \beta_s = 1, \alpha_s = 1, -3, \phi = 0.9, \sigma_{ss} = (0.035)^2, \sigma_{\eta\eta} = (0.001)^2, \) and \( \rho_{\eta\eta} = 0.16 \). When \( \alpha_s = 1 \) the FRUH is true and when

<table>
<thead>
<tr>
<th>Table 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monte Carlo estimates of bias in levels regressions ( f = s_t + u_{pt}, \ u_{pt} = 0.9u_{pt-1} + \eta_{s_t} s_t + u_{st} )</td>
</tr>
<tr>
<td>( u_{st} = \alpha_s u_{pt-1} + \eta_{st} = idN \left( \begin{array}{c} \mu_s \ \eta_{st} \end{array} \right) \right) )</td>
</tr>
<tr>
<td>Estimated regression: ( f = a + bs + e^a )</td>
</tr>
<tr>
<td>( T=100 )</td>
</tr>
<tr>
<td>( \alpha_s = 1 )</td>
</tr>
<tr>
<td>( \alpha_s = -3 )</td>
</tr>
</tbody>
</table>

\( \alpha_s = 1 \)

-0.000

-0.005

1.000

-0.079

0.000

0.023

1.000

-0.122

(0.642)

(0.605)

(0.653)

(0.629)

\( \alpha_s = -3 \)

-0.000

-0.018

1.000

-0.285

-0.000

-0.065

1.000

-0.303

(0.644)

(0.595)

(0.644)

(0.622)

\(^{15}\) The Stock–Watson DOLS/DGLS and Johansen ML estimators of \( \beta_s \), asymptotically remove the effects of \( \theta \) and \( \Delta_{\eta\eta} \) and so are asymptotically unbiased and efficient.

\(^{16}\) The parameters for the Monte Carlo experiment were calibrated from monthly data on UK spot and forward rates quoted in US dollars. To check whether the Monte Carlo results were sensitive to the normality of the errors we also computed results based on a bootstrap approach using the US/UK exchange rates. The bootstrap results were essentially identical to the Monte Carlo results and are therefore omitted. Details are available upon request.
\[ \alpha_s = -3 \] it is not. In both cases the forward premium is highly autocorrelated. Table 12 gives the results of OLS applied to the levels regression Eq. (12a) for samples of size \( T = 100 \) and \( T = 250 \). In both cases the magnitude of the OLS bias in the coefficients is negligible but the OLS standard errors are quite biased which cause substantial size distortions in the nominal 5% \( t \)-tests of the hypothesis that \( \beta_s = 1 \) and \( \mu_c = 0 \).

The derived triangular representation Eqs. (14a), (14b), (14c) and (14d) for \((s_{t+1}, f_t)'\) can be used to show that the OLS estimate of \( \beta_f \) in the levels regression Eq. (3) will be asymptotically downward biased even if \( \alpha_s = 1 \) (the FRUH is true). To illustrate, consider the VECM Eqs. (11a) and (11b) 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 Eqs. (14a), (14b), (14c) and (14d) simplifies to

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

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

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

so that \( \theta = \sigma_{sf}/\sigma_{ff} \) and \( \Delta_0 = 0 \). 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 Eq. (3) will suffer from bias even if the FRUH is true. This result is due to the fact that the long-run covariance matrix of \( v_{s,t+1} \) is not diagonal. To illustrate the magnitude of the bias, Table 13 reports OLS estimates of the levels regression Eq. (3) when data are generated from Eqs. (11a) and (11b). The OLS estimate of \( \beta_f \) 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% \( t \)-test rejects the null that \( \beta_f = 1 \) about 30% of the time when the null is true. Table 14 also summarizes results for the Stock–Watson DOLS estimator. In all cases, the Stock–Watson estimator is

<table>
<thead>
<tr>
<th>( T=100 )</th>
<th>( T=250 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_s = 1 )</td>
<td>( \alpha_s = -3 )</td>
</tr>
<tr>
<td>( a )</td>
<td>( t_{p=0} )</td>
</tr>
<tr>
<td>( -0.000 )</td>
<td>(-0.030 )</td>
</tr>
<tr>
<td>( (0.275) )</td>
<td>( (0.293) )</td>
</tr>
<tr>
<td>( -0.000 )</td>
<td>(-0.007 )</td>
</tr>
<tr>
<td>( (0.309) )</td>
<td>( (0.278) )</td>
</tr>
</tbody>
</table>

* See notes for Table 12.
Table 14
Monte Carlo estimates of bias in levels regressions. Estimated regression: \( s_{t+1} = a + b f_t + \sum_{k=-3}^{3} \gamma_k \Delta f_{t-k} + e_{t+1} \)

<table>
<thead>
<tr>
<th>( T=100 )</th>
<th>( T=250 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_z=1 )</td>
<td>( a )</td>
</tr>
<tr>
<td>-0.000</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0.171)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>( \alpha_z=-3 )</td>
<td>-0.000</td>
</tr>
<tr>
<td>(0.812)</td>
<td>(0.142)</td>
</tr>
</tbody>
</table>

\(^a\) See notes for Table 12.

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 are moderate size distortions in the nominal 5% \( t \)-tests of \( \beta_f=1 \) and \( \mu=0 \) for \( T=100 \) but the distortions dissipates as \( T \) increases.

References


Barnhart, S.W., McNown, R., Wallace, M.S., 1998. Some answers to puzzles in testing efficiency of the foreign exchange market. Unpublished manuscript, Department of Economics, Clemson University, Clemson, SC.

Barkoulas, J., Baum, C.F., 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.


