Testing Uncovered Interest Parity: A Continuous Time Approach\textsuperscript{1}

Antonio Diez de los Rios\textsuperscript{2}  
CEMFI and Universidad de Malaga  
adriosg@cemfi.es

Angel Leon  
Universidad de Alicante  
aleon@ua.es

Enrique Sentana  
CEMFI, CEPR and LSE-FMG  
sentana@cemfi.es

October 2003  
Preliminary Version. Please, Do Not Quote

\textsuperscript{1}We would like to thank Manuel Arellano, Lars Hansen, Robert Hodrick, Nour Meddahi as well as participants at the European Meeting of the Econometric Society (Stockholm, 2003) for their useful comments and suggestions, though we remain responsible for any remaining errors.

\textsuperscript{2}Corresponding author: CEMFI, Casado del Alisal, 5, E-28014, Madrid, Spain, Phone: (+34) 914 29 05 51, Fax: (+34) 914 29 10 56, E-mail: adriosg@cemfi.es
Abstract

This paper investigates whether the issue that exchange rates evolve on a much finer time-scale than the frequency of observations has any impact on traditional uncovered interest parity (UIP) tests. To this end, we specify a multivariate continuous-time model for testing UIP. Derivation of the conditions on the parameters of the continuous-time system that guarantee UIP are provided using the Wold decomposition and exploiting the first order companion VAR form of the model. It is shown that, contrary to the (discrete-time) VAR approach, these conditions do not depend on the sampling mechanism of the data set, nor the test is distorted by the temporal aggregation problem. Furthermore, our approach generates asymptotically efficient estimators of the parameters of the continuous-time system.
1 Motivation

During the last twenty five years, the majority of studies have rejected the hypothesis of uncovered interest parity, that is, that the (nominal) expected return to speculation in the forward foreign exchange market conditioned on available information is zero. These studies have typically regressed ex post rates of depreciation on a constant and the forward premium, and have rejected the null hypothesis that the slope coefficient is one. In particular, a robust result in these studies is that this slope is negative, and significantly different from one. This phenomenon is known as the “forward premium puzzle” and implies that, contrary to theory, high domestic interest rates relative to those in the foreign country predict a future appreciation of the home currency.

Different explanations have been given to this issue, but none of them is fully satisfactory. For instance, rejection of the hypothesis of uncovered interest parity (UIP from now on) can be related to the existence of a rational risk premium in the foreign exchange rate market. However, the literature has shown a failure of these premises at the same level than the rejection of the UIP. As a good example, Backus, Foresi and Telmer (2001) characterize the anomaly in the context of affine models of the term structure. They find that the forward premium puzzle requires that either the state variable have asymmetric effects on state prices in different currencies in a way that “[to use a concrete example] U.S money growth has a larger influence than British monetary policy on US$ interest rates, but a smaller influence on the US$ pricing kernel”, or that negative realizations of the nominal interest rate can occur with a positive probability. A second line of research have claimed the existence of a “peso problem”, or even have abandoned the “rationality assumptions” to reconcile theory and the puzzle.¹

Nonetheless, the claim that we present here is different as long as we are interested in whether the traditional studies on uncovered interest are testing the right things, and whether the estimators of the parameters have good properties. In this line, it is interesting to emphasize that evidence against UIP has been lessened in recent studies. Flood and Rose (2001) find that UIP works better in the 1990’s, while Baillie and

¹See Lewis (1989) for details of the “peso problem approach”, and Mark and Wu (1998) for a model that adapts the overlapping-generation noise-trader model of De Long et al. (1990)
Bollerslev (2000) and Bekaert and Hodrick (2001) find that evidence against uncovered interest parity is much less strong under finite sample inference than under traditional asymptotic theory.

Particularly, there are two main approaches to test UIP. The first one has been previously mentioned and it is a regression approach where we consider the estimation of a single equation that relates the increment of the spot exchange rate over the contract period to the forward premia. In this setup we only need to care that when the contract period of the forward contract is longer than the sampling interval there is an overlapping problem that induces serial correlation into the regression error. In this line, Hansen and Hodrick (1980) show how to use Hansen’s Generalized Method of Moments (GMM) to obtain standard errors that are robust to autocorrelation. On the other hand, we can specify the joint process that is driving the forward premia and the increment on the spot exchange rate over the sampling period and test the constraints that UIP implies for the dynamic evolution of both variables. This way we can exploit the time series properties of the data to achieve efficiency gains in the test (Hallwood and MacDonald, 1994).

In this paper we focus on the impact of the temporal aggregation problem on the statistical properties of these two traditional tests of UIP, where by temporal aggregation we mean the fact that exchange rates evolve on a much finer time-scale than the frequency of observations. While in a lot of economic examples the sampling frequency is given because collecting data is expensive in terms of time and money (i.e. output, labor force statistics), this is not the case of financial data any more. Here, and in particular for exchange and interest rates, the data sampling frequency is an election to the researcher. The insight why this is an important problem that we need to be aware comes from the result in Hansen and Hodrick (1980) that we can achieve gains in the asymptotic power of tests by means of using overlapping data. As an example of the impact of the sampling mechanism on power issues recall that on the early work on UIP, as in Frenkel (1977), data where sampled in such a way to produce a data set with non-overlapping residuals with the corresponding waste of degrees of freedom and its implication in terms of power.
Our claim is that if the number of observations per contract period is a large number (which in terms of power of tests should be a good thing), the degree of overlapping may completely distort the standard errors estimates and render that standard GMM asymptotic theory is no longer a good approximation to the finite sample distribution of UIP tests. As an example think that we are interested in testing UIP using 3-month forward contracts and weekly data as in Hansen and Hodrick (1980). Then, there is an overlapping degree of 12 periods and it seems that standard GMM asymptotics are likely to work well. But if we want to use daily data (five days per week), we will have an overlapping degree of 60 periods that is probably distorting inference. Moreover, this problem can be worsened for longer contracts, or if we want to employ intra-daily data for testing purposes to achieve power improvements.

In this situation we can only rely on estimating a joint process, generally a vector autoregression, for the forward premia and the spot exchange rate. Thus, the difference operation on the spot exchange rate is taken over the sampling interval and therefore there is no overlapping in the residuals of the joint process and there is no need to make serial correlation adjustments in the estimation of the standard errors. However, the election of the sampling interval has also implications for the joint process of our variables. To illustrate this problem assume that the forward premia and the rate of depreciation are generated twice a day by a first order vector autoregressive process VAR(1), but we only use the end-of-day value of both variables. In this setup, it can be shown that our observations will satisfy a vector autoregressive moving average VARMA model.

Moreover, having a model that is “sampling-frequency-proof” helps us to reduce any misspecification problem that may arise from equating the data generating interval with the sampling interval when the former is in fact finer. This is what is called the temporal aggregation problem and is another important aspect from our framework because having a misspecified model can be translated into an inconsistent UIP test.

Motivated by these issues, we reconsider tests for uncovered interest rate parity using a continuous time approach. In particular, we assume that there is an underlying continuous time process that is only observed at discrete points of time and we estimate
the parameters of this unobservable process. This approach has the advantage that we can accommodate situations with a large ratio of observation periods per contract periods (with the corresponding gains in terms of asymptotic power) and at the same time have a framework where the model that the observed data follow is the same regardless the sampling frequency. Thus, the key point in the paper is acknowledging that the dynamic of economic variables evolve through time even when we are not sampling them.

In particular, we derive the conditions to test uncovered interest parity in a continuous-time model using the Wold decomposition. We also show that when the model of interest is a continuous-time VAR(1) process we can exploit its discrete time first-order companion VAR representation form to obtain the conditions that guarantees UIP as well as for estimation purposes. In this situation, we can evaluate the likelihood of the model via the prediction error decomposition using Kalman filtering techniques, and therefore we can obtain asymptotically efficient estimators of the parameters of the continuous time model (provided that the model accurately describes the law of motion for the discrete time data).

The paper is organized as follows. Section 2 presents the basic testing equations that have been employed in the literature on uncovered interest rate parity. In section 3, we discuss these tests. Section 4 details the continuous-time framework and the conditions for uncovered interest rate in continuous time models. An example is presented in section 5. Final remarks and future lines of research are presented in section 6.

2 Uncovered Interest Parity: Basic Testing Equations.

Empirical work on the relation between the forward premium in the foreign exchange market and the expected change in the spot exchange rate has been an area of active research for the last twenty-five years. In particular, an important building block of this relationship has been the Uncovered Interest Parity, which states that the (nominal)

\footnote{For a complete review of the literature, see Hodrick (1987), Engel (1996) and Wang and Jones (2002).}
expected return to speculation in the forward foreign exchange market conditioned on available information is zero:

\[ E_t [s_{t+\tau} - s_t] = f_t - s_t \]  

(1)

where \( s_t \) is the logarithm of the spot exchange rate \( S_t \), \( f_t \) is the logarithm of the forward rate \( F_t \) contracted at \( t \) and matures at \( t + \tau \). As a consequence, the (log) forward exchange rate is an unbiased predictor of the \( \tau \)-periods ahead (log) spot exchange rate. For this reason, UIP is also known as “Unbiasedness Hypothesis”.

Although a main criticism made to UIP is that it pays no attention to issues of risk aversion and intertemporal allocation of wealth, Hansen and Hodrick (1983) have shown that equation (1) with an additional constant term is consistent with a model of rational maximizing behaviour in which assets are priced by a no arbitrage restriction. In these economies the (dollar) price \( P_t \) of a claim to the future and uncertain cash flow of \( D_{t+1} \) (dollars) one period later must satisfy:

\[ P_t = E_t [D_{t+1}M_{t+1}] \]  

(2)

where \( M_{t+1} \) denotes the stochastic discount factor (SDF). In particular, (2) can be understood as the Euler equation of an agent intertemporal maximization problem, while the SDF can be related to the agents’ nominal intertemporal marginal rate of substitution\(^3\). Therefore, given that a position in a one period contract in the forward foreign exchange market involves no payment at date \( t \) while a profit of \( F_t - S_{t+1} \), the pricing relationship (2) implies that

\[ 0 = E_t [(F_t - S_{t+1})M_{t+1}] \]

Dividing by \( S_t \) and rearranging we get

\[ \frac{1}{S_t} E_t [S_{t+1}M_{t+1}] = \frac{F_t}{S_t} E_t [M_{t+1}] \]

When the exchange rates and the SDF are log-normally distributed this equation can be written as

\[ E_t [s_{t+1} - s_t] = -\frac{1}{2} Var_t [s_{t+1}] - Cov_t [s_{t+1}, m_{t+1}] + f_t - s_t \]

\(^3\)See Cochrane (2000) for an extended discussion of the stochastic discount factor.
where $m_{t+1}$ is the logarithm of the SDF, while the conditional variance and covariances are denoted by $\text{Var}_t[\cdot]$ and $\text{Cov}_t[\cdot]$. Moreover, under the additional assumption of constancy of conditional second moments, we get

$$E_t[s_{t+1} - s_t] = \alpha + f_t - s_t$$

where $\alpha = -\frac{1}{2}\text{Var}_t[s_{t+1}] - \text{Cov}_t[s_{t+1}, m_{t+1}]$ is a constant. This proposition is known as “Modified Unbiasedness Hypothesis”, and it is the hypothesis we focus on this paper as Uncovered Interest Parity.

### 2.1 The Limited Information Approach

Ordinary least squares tests of the “Modified Unbiasedness Hypothesis” can be easily derived. To this end, assume rational expectations (RE) in foreign exchange markets:

$$s_{t+\tau} - s_t = E_t[s_{t+\tau} - s_t] + \mu_{t+\tau, \tau}$$

where $\mu_{t+\tau, \tau}$ is a rational expectation error with zero mean and uncorrelated with any variable in the time $t$ information set.

Combining eqs (3) and (4) makes:

$$s_{t+\tau} - s_t = \alpha + (f_t - s_t) + \mu_{t+\tau, \tau}$$

which has motivated the following OLS regression, as in Geweke and Feige (1979), as the usual starting point:

$$s_{t+\tau} - s_t = \lambda_0 + \lambda_1 (f_t - s_t) + \mu_{t+\tau}$$

to test UIP. “Unbiasedness Proposition” implies that $\lambda_0 = 0$ and $\lambda_1 = 1$, while in the case of the “Modified Unbiasedness Proposition”, we just need that $\lambda_1 = 1$. However, the RE assumption implies that errors are not autocorrelated as long the sampling interval is equal or larger than $\tau$. For this reason, different authors, as Frenkel (1977) among others, have sampled data to produce a data set with non-overlapping residuals with the corresponding waste of degrees of freedom.

On the other hand, Hansen and Hodrick (1980) show how to use overlapping data in order to increase to the sample size, which will be reflected in corresponding gains in
the asymptotic power of tests. Using Hansen’s Generalized Method of Moments, they obtain asymptotic standard errors that take into account the serial correlation induced into the regression error when the prediction horizon is higher than the sampling interval of the data. In the same way, standard errors robust to conditional heteroskedasticity can also be computed.

Nonetheless, some points with respect to the single equation approach have to be emphasized. Firstly, the estimate of the asymptotic covariance matrix is very sensitive to the election of the bandwidth and the kernel chosen in its estimation. This results in inferences that are severely distorted and over-rejection in finite samples (den Haan and Loven, 1996). Secondly, the construction of the estimate of the covariance matrix is infeasible if we want to use high-frequency data in order to test UIP, given that the degree of overlapping tends to infinity, and therefore, standard GMM asymptotic theory does no longer apply.

2.2 The Full Information Approach

Up to this point we have considered the estimation of a single equation that relates the $\tau$-increment in the spot exchange rate, $s_{t+\tau} - s_t$, to the forward premia $p_t$. A second approach considers estimating, usually by maximum likelihood (ML), a joint covariance stationary and multivariate normal process for the (demeaned) first difference of the spot exchange rate $\Delta s_t$, and the forward premia $p_t$. In this approach the difference operation on the spot exchange rate is taken over the sampling interval, and, consequently, the UIP condition is expressed as:

$$E_t \left[ \sum_{i=1}^{\tau} \Delta s_{t+i} \right] = p_t$$

(6)

This hypothesis is tested imposing constraints on the joint process that describes both series, and to this end, we exploit the Wold decomposition to obtain the projections of each one of the elements in the sum with respect to the information set defined by $\{\Delta s_t, p_t, \Delta s_{t-1}, p_{t-1}, \ldots\}$.

$^4$This procedure was first discussed in Hansen and Hodrick (1980), although the efficient maximum likelihood estimator was not used there due to computational infeasibility.
To illustrate this technique, let $y_t = [\Delta s_t, p_t]'$ be a wide-sense stationary stochastic process whose law of motion is described by the following vector moving average representation (Wold decomposition):

$$y_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$$  \hspace{1cm} (7)

where $\varepsilon_t$ is a 2-dimensional vector of white noise disturbances and contemporaneous covariance matrix $E(\varepsilon_t \varepsilon_t') = \Sigma$. Moreover, $C(L)$ is a $2 \times 2$ matrix of polynomials in the lag operator such that the sequence of matrices $C \equiv \{C_j : j \in J\}$ being $J$ the set of all integers must be “square summable”, that is $tr\left[\sum_{j=0}^{\infty} C_j \Sigma C_j'\right] < \infty$.

Given this vector moving average representation, we can just apply the Wiener-Kolmogorov formulae for linear least squares prediction to obtain $E_t\Delta s_{t+i}$ for $i = 1, ..., \tau$

$$E_t\Delta s_{t+i} = C_{21}(L)\varepsilon_{1t} + C_{22}(L)\varepsilon_{2t} \hspace{1cm} i = 1, ..., \tau$$

where $C_{ij}(L)$ is the $ij$-element of $C(L)$ and $[.]_+$ is the annihilation operator and means “ignore the negative powers of $L$”. Therefore, once we aggregate the corresponding set of projections, the LHS of (6) can be expressed as:

$$E_t \left[ \sum_{i=1}^{\tau} \Delta s_{t+i} \right] = \sum_{i=1}^{\tau} \left[ \frac{C_{21}(L)}{L^i} \right] \varepsilon_{1t} + \sum_{i=1}^{\tau} \left[ \frac{C_{22}(L)}{L^i} \right] \varepsilon_{2t} \hspace{1cm} (8)$$

Noting that the Wold decompositon implies that

$$p_t = C_{11}(L)\varepsilon_{1t} + C_{12}(L)\varepsilon_{2t} \hspace{1cm} (9)$$

it is straightforward to see that (8) and (9) are equal for every realization of $\varepsilon_t$, if and only if the following two conditions are fulfilled:

$$C_{11}(L) = \sum_{i=1}^{\tau} \left[ \frac{C_{21}(L)}{L^i} \right]_+$$

$$C_{12}(L) = \sum_{i=1}^{\tau} \left[ \frac{C_{22}(L)}{L^i} \right]_+$$

Note that in this model the (log) spot and forward exchange rates are cointegrated, given that its difference, the forward premia, is a stationary process. Moreover, if both the forward premia and the first difference of the spot exchange rate follow a vector
autorregression process (VAR) the model will have a resemblance to the Campbell and Shiller (1987) set-up for testing present value models.

However, there is another way to proceed that has been proved to be more useful in applied work and consists in the estimation of a finite order vector autoregression (VAR) representation with a sufficient number of lags as a good approximation of the infinite order Wold decomposition of the difference of the spot and forward exchange rates. To show this, consider that $y_t = [\Delta s_t, p_t]'$ follows a VAR($p$) model:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t$$

where, then again, $\varepsilon_t$ is a 2-dimensional vector of white noise disturbances and contemporaneous covariance matrix $E(\varepsilon_t\varepsilon_t') = \Sigma$.

Exploiting the first order companion form of this model:

$$
\begin{pmatrix}
  y_t \\
  y_{t-1} \\
  \vdots \\
  y_{t-p+1}
\end{pmatrix}
= 
\begin{pmatrix}
  A_1 & A_2 & \cdots & A_p \\
  I & 0 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
  y_{t-1} \\
  y_{t-2} \\
  \vdots \\
  y_{t-p}
\end{pmatrix}
+ 
\begin{pmatrix}
  \varepsilon_t \\
  0 \\
  \vdots \\
  0
\end{pmatrix}
$$

we can obtain the forecast of $x_{t+i}$ for $i = 1, ..., \tau$ based on the information set defined by this VAR model just using the law of iterated expectations:

$$E_t x_{t+i} = \Theta^i x_t$$

If we define an indicator vector $e_j$ with the same dimension that $x_t$ and that have a one in the $j$-th position, and zero in the other one, it is clear that we can express the projection of $\Delta s_{t+i}$ conditioned on the information set, $E_t \Delta s_{t+i}$ as:

$$E_t \Delta s_{t+i} = e_2' \Theta^i x_t$$

Therefore, the LHS of (6) can be expressed as:

$$E_t \left[ \sum_{i=1}^{\tau} \Delta s_{t+i} \right] = e_2' \left[ \sum_{i=1}^{\tau} \Theta^i \right] x_t = e_2' (\Theta^\tau + \Theta^{\tau-1} + ... + \Theta) x_t$$

while the right hand side (RHS) is:

$$p_t = e_1' x_t$$
Therefore, the vector of restrictions on the VAR parameters from the UIP for a τ-period forward contract may be written as:

\[ e_2'(\Theta^\tau + \Theta^{\tau-1} + \ldots + \Theta) = e_1' \]

\[ e_2' \left( \sum_{i=1}^{\tau} \Theta^i \right) = e_1' \]

Given that the difference operation is made over the sampling interval, there is no overlapping in the regression residuals of the VAR model and, for that reason, no serial correlation adjustment is required in the estimation of the standard errors of this model.

Moreover, in this paper we extend this approach to cover also the UIP restrictions on VARMA models. In particular, if \( y_t = [\Delta s_t, p_t]' \) follows a VARMA\((p,q)\) model

\[ y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t + C_1 \varepsilon_{t-1} + \ldots + C_q \varepsilon_{t-q} \]

all that we need to note is that the results derived before applies directly to the first order companion form of this VARMA model:

\[
\begin{pmatrix}
  y_t \\
y_{t-1} \\
  \vdots \\
y_{t-p+1} \\
  \varepsilon_t \\
  \varepsilon_{t-1} \\
  \vdots \\
  \varepsilon_{t-q+1}
\end{pmatrix}
= 
\begin{pmatrix}
  A_1 & A_2 & \cdots & A_p & C_1 & C_2 & \cdots & C_q \\
  I & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
  0 & 0 & \cdots & 0 & I & 0 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
  y_{t-1} \\
y_{t-2} \\
  \vdots \\
y_{t-p} \\
  \varepsilon_{t-1} \\
  \varepsilon_{t-2} \\
  \vdots \\
  \varepsilon_{t-q}
\end{pmatrix}
+ 
\begin{pmatrix}
  \varepsilon_t \\
  \vdots \\
  \varepsilon_t \\
  \varepsilon_t \\
  \varepsilon_t \\
  \vdots \\
  \varepsilon_t \\
  \varepsilon_t
\end{pmatrix}
\]

\[ x_t = \Theta x_{t-1} + u_t \]

3 Understanding Uncovered Interest Parity Tests

In this paper we are interested in whether the traditional tests on UIP are doing their job, and whether the estimators of the parameters have good properties. In this line, it can be shown that the full information approach is more efficient than the (limited information) regression procedure given that the time series properties of the data are exploited (Hallwood and MacDonald, 1994). However, what is not usually emphasized is that this efficiency gain is not a “free lunch”, given that full information tests rely on the validity of the auxiliary model. To be precise, if the auxiliary model is not a good
representation of the Wold decomposition (i.e. wrong number of lags selection in a VAR process), we lose the consistency property of the tests based on the regression approach, and therefore, it is possible to reject the UIP hypothesis when it is the true model.

Moreover, the VAR(MA) approach does not have a time invariance property in the sense that if data generating interval is finer than the sampling interval then the model is not independent from the sampling frequency. That is if the same series were sampled more frequently then the model that these observations follow is not going to be the same. Given that financial data and exchange rates evolve on a much finer time-scale than the frequency of observations, it is important to have a model that is robust to this issue. This is because choosing the wrong model has the disastrous consequence of having parameter estimates distorted by temporal aggregation biases, and losing the consistency property of the test. However, before trying to solve this issue, it seems worth it to expend some time analyzing these two problems in depth.

3.1 Efficiency versus Consistency

Implied regression  To address the issue of efficiency and consistency in UIP tests, we first introduce the concept of implied regression. To do so, let

\[ x_{t+1} = \Theta x_t + u_{t+1} \]  

be the corresponding first-order VAR form of the system where \( x_t \) defines the elements in the time \( t \) information set implied by the VAR(MA) model. Recalling that \( s_{t+\tau} - s_t = \Delta s_{t+\tau} = \sum_{i=1}^{\tau} \Delta s_{t+i} \) we can just substitute (11) into this expression to get the regression that we are implicitly running when testing UIP through a VAR(MA) model

\[
\begin{align*}
\Delta s_{t+\tau} &= e_2' \left[ \sum_{i=1}^{\tau} \Theta^i \right] x_t + e_{2j}' \left[ \sum_{j=1}^{\tau} \sum_{k=0}^{j-1} A^k \varepsilon_{t+j} \right] \\
\Delta s_{t+\tau} &= \pi' x_t + u_{t+\tau}
\end{align*}
\]

where \( \pi \) is a vector containing the implied coefficients of the projection of the \( \tau \)-increment of the spot exchange rate onto the VAR(MA) implied time \( t \) information set, and where \( u_{t+\tau} \) is a noise term that has a moving average (MA) structure of order \( \tau - 1 \).

This expression is very useful to understand what is the object that we are testing when estimating VAR(MA) models. The first note of caution that comes from (12)
is that \( \pi' = e_2' [\sum_{i=1}^{r} \Theta^i] \) do not converge to the true projections of \( \Delta_t s_{t+\tau} \) onto \( x_t \) unless the model is correctly specified. Thus rejection that \( e_2' [\sum_{i=1}^{r} \Theta^i] = e_1' \) does not necessarily reflect that UIP does not hold. This fact makes specially important the implementation of specification tests.

A second remark is that we cannot careless compare both the traditional regression and the VAR(MA) approach because the latter is implicitly running a different regression. Therefore both test are equivalent under the null hypothesis that UIP holds, but on the other side they have different alternative hypotheses. As a direct consequence, if we want to properly compare both approaches we have to run the regression \( \Delta_t s_{t+\tau} \) on \( x_t \), that is, the variables that are in the time \( t \) information set defined by the VAR(MA) model

**Implied slope coefficient**  Another statistic that can be employed for comparison purposes between both approaches is the implied slope coefficient. This statistic, among others, has been extensively used to examine (long horizon) predictability of currency excess returns in Bekaert and Hodrick (1992) and Bekaert (1995).

The idea behind this statistic is fairly simple as it only consists in getting the implied coefficient of a regression of the \( \tau \)-increment of the spot exchange rate against the forward premia in terms of the parameters of the VAR(MA) model. In particular, the implied slope coefficient can be calculated as

\[
\beta = \frac{e_2' [\sum_{i=1}^{r} \Theta^i] \Psi e_1}{e_1' \Psi e_1}
\]

where \( \Psi = E [x_t x'_t] \) is the unconditional variance of \( x_t \) implied by the VAR(MA) model. Note that the numerator is just the implied covariance of the \( \tau \)-increment of the spot exchange rate with the forward premia, \( \text{Cov} [\Delta_t s_{t+\tau}, p_t] \), while the denominator is just the implied variance of the forward premia \( \text{Var}[p_t] \). Then again, \( \beta \) in (13) only converges to the true projection coefficient of \( \Delta_t s_{t+\tau} \) onto \( p_t \) when the VAR(MA) model is the true data generating process.

Even more, (13) can be used to emphasize that the traditional regression and the VAR(MA) approach are testing different objects. To illustrate this, let \( \tau = 1 \) and assume
that the data generating process for \( y_t = [p_t, \Delta s_t]' \) be a VAR(1) model, so

\[ y_{t+1} = Ay_t + \varepsilon_{t+1} \]

Then it is clear that UIP holds in this setup when

\[ e_2' A = e_1' \implies a_{21} = 1, \quad a_{22} = 0 \]

where \( a_{ij} \) is the \( ij \)-element of \( A \).

This very simple model brings into light that testing \( \beta = 1 \) does not guarantee that the conditions for UIP are fulfilled. To show this, recall that if the implied beta (which coincides exactly with the true slope coefficient) is equal to one then

\[ \beta = \frac{e_2' A \Psi e_1}{e_1' \Psi e_1} = 1 \]

Rearranging

\[ a_{21} = 1 - a_{22} \frac{\psi_{21}}{\psi_{11}} \]

where \( \psi_{ij} \) is the \( ij \)-element of \( \Psi \). This expression gives you the set of values for \( a_{21} \) in terms of \( a_{22} \) that makes that \( \beta = 1 \). However, note that \( \beta = 1 \) makes that UIP holds only when \( a_{22} = 0 \). This is because \( E_t [s_{t+\tau} - s_t] = p_t \) implies that the expected return to speculation in the forward foreign exchange market conditioned on all the available information is zero, which is a stronger assumption than having that the expected return conditioned on \( p_t \) is zero.

### 3.2 The temporal aggregation problem

A second problem that we should be aware of is related to the fact that if the data generating interval is finer than the sampling interval then the data generating process and the data observed process are not the same. This problem is known in the literature as the “temporal aggregation problem” and it will cause the estimates of the discrete time model to be distorted by spurious Granger causality, serial persistence in residuals...

But even more, when we ignore this problem then the model is likely to be misspecified rendering the UIP test inconsistent.
Just for illustration purposes, assume that $y_t^{(1)} = [p_t, \Delta s_t]'$, that is, the forward premia and the increment of the spot exchange rate, are generated twice a day by a VAR(1) process, but we only observe the end-of-day value of both variables.

$$y_{t+1}^{(1)} = Ay_t^{(1)} + \varepsilon_{t+1}$$

where $E[\varepsilon_t\varepsilon_t'] = \Sigma$.

Then the interesting question consist in finding out the process that follows the variable that we observe, that is $y_t^{(2)} = [p_t, \Delta_2 s_t]'$. To this end, we just need to note that

$$\Delta_2 s_{t+2} = \Delta s_{t+1} + \Delta s_t$$

Substituting the process for $y_t^{(1)}$ and after some algebraic manipulations we get that:

$$\left( \begin{array}{c} p_{t+2} \\ \Delta_2 s_{t+2} \end{array} \right) = A^2 \left( \begin{array}{c} p_t \\ \Delta_2 s_t \end{array} \right) + \varepsilon_{t+2} + \left[ \left( \begin{array}{c} 0 \\ \varepsilon_2' \end{array} \right) + A \right] \varepsilon_{t+1} + \left( \begin{array}{c} 0 \\ \varepsilon_2' \end{array} \right) A \varepsilon_t$$

which implies that $y_t^{(2)}$ follows a VARMA(1,1) process. This is clearer once we note that we can rewrite this process as

$$\left( \begin{array}{c} p_{t+2} \\ \Delta_2 s_{t+2} \end{array} \right) = A^2 \left( \begin{array}{c} p_t \\ \Delta_2 s_t \end{array} \right) + v_{t+2} + Cv_t$$

where $v_t = y_t^{(2)} - E_t - 2y_t^{(2)}$ is the “one-day”-ahead prediction error with

$$E[v_tv_t'] = \Omega_v = \Sigma + \left[ \left( \begin{array}{c} 0 \\ \varepsilon_2' \end{array} \right) + A \right] \Sigma \left[ \left( \begin{array}{c} 0 \\ \varepsilon_2' \end{array} \right) + A \right]'$$

and

$$C = \varepsilon_2' A \Sigma \Omega_v^{-1}$$

The usual way to circumvent the temporal aggregation problem is to assume that there is an underlying and unobservable continuous time process on which we only have observations at discrete points of time. In particular, this is what is done in the next section.

Moreover, VAR(MA)-based tests of UIP have the unattractive feature that they are not robust to the election of the sampling mechanism. In particular, if we acknowledge that data is sampled each $\delta \leq \tau$ periods, we have the following UIP restrictions
\[ e'_2(\Theta^{r/\delta} + \Theta^{r/\delta - 1} + \ldots + \Theta) = e'_1 \]

This property is an important drawback, given that the restrictions on the coefficients that imply UIP depends on the sampling interval \( \delta \), which is an election of the researcher, and therefore something arbitrary. Furthermore, if we want to interpret UIP test as the result of an asset pricing model, as in Hansen and Hodrick (1983), this fact is specially important as long as the sampling period is a variable that appears in the econometric test but it does not in the asset pricing model. This problem is usually obscured because \( \delta \) usually is equal to 1. In the next section we show that this sampling interval dependence is removed when we specify the model in continuous time.

4 A continuous time framework

In this section we provide the conditions that guarantee UIP in continuous time models. For this reason, we make use of the framework of Phillips (1991) and Chambers (2001) to state a model for the spot and forward exchange rate in continuous time in which there is a cointegration relationship between these two variables. In particular, let the model of interest for the (demeaned) forward premia and the spot exchange rate take the following form:

\[
p(t, \tau) = u_1(t) \tag{14}
\]

\[
ds(t) = u_2(t)dt \tag{15}
\]

where, \( p(t, \tau) \equiv f(t, \tau) - s(t) \) is the forward premia, being \( S(t) \) the spot exchange rate and \( F(t, \tau) \) the \( \tau \) day forward exchange rate, respectively, and \( s(t) = \ln S(t), f(t, \tau) = \ln F(t, \tau) \). Finally, \( u(t) = [u_1(t), u_2(t)]' \) is a stationary continuous time residual whose Wold decomposition is defined through the following convolution as in Hansen and Sargent (1991):

\[
u(t) = \int_0^\infty \phi(h)dW_n(t - h)
\]
where the Wiener processes $dW_a(t) = [dW_{a1}(t), dW_{a2}(t)]'$ have the following covariance structure:

$$ E[dW_a(t)dW_a(t)'] = \Sigma dt $$

and being $\phi(h)$ “square integrable”, that is, $\text{tr} \left[ \int_0^\infty \phi(h)\Sigma_u\phi(h)'dh \right] < \infty$

Note that this is the continuous-time counterpart of (7), where the integral plays the role of the sum, $\phi(h)$ plays the role of $C_j$, while $dW_a(t-h)$ is related to $\varepsilon_{t-j}$. Even more, (14) embodies a cointegration relationship between the spot and the forward exchange rate, given that its difference is a stationary process and therefore both spot and forward exchange rates cannot diverge in the long run. In particular, this relationship is established by the covered interest rate parity, and as a direct consequence, $u_1(t)$ is interpreted as the $\tau$ day interest rate differential between home and foreign country.

On the other hand note that solving (15) we obtain:

$$ s(t) = \int_0^t u_2(r)dr + s(0) $$

so that the first term of the (log) spot exchange rate can be regarded as the outcome of accumulated innovations over the interval $[0,t]$ which resembles a I(1) variable in discrete-time.

However, note that (14) implies that $Ds(t) = u_2(t)$ where $D \equiv \frac{d}{dt}$ is the mean square differential operator. Therefore, we are implicitly assuming that the sample paths for the spot exchange rate $s(t)$ are differentiable and therefore that the infinitesimal change in $s(t)$ is smooth. This is an assumption that does not seem to be a good description of the behaviour of exchange rates. Then, we solve this issue by adding a Wiener process (which is not differentiable) to (14):

$$ p(t, \tau) = u_1(t) $$

$$ ds(t) = u_2(t)dt + dW_s(t) $$

where the Wiener processes $dW(t) = [dW_{a1}(t), dW_{a2}(t), dW_s(t)]'$ have the following covariance structure:

$$ E[dW(t)dW(t)'] = \Sigma dt $$

Note that by formal differentiation this augmented model collapses to (14) and (15) when $dW_s(t) = 0 \forall t$. 

16
4.1 Conditions for Instantaneous Uncovered Interest Parity

Given that we have a continuous time model, it is natural as a first step to look for the conditions that make uncovered interest parity hold in a framework where agents can invest in instantaneous forward contracts. To this end, let the forward exchange rate contract period $\tau$ tends to zero, which means that $f(t, 0)$ is the (log) instantaneous forward exchange rate. Then, it is clear that instantaneous UIP holds when:

$$E_t[ds(t)] = p(t, 0)dt = u_1(t)dt$$

that is, when the expected infinitesimal change in the spot exchange rate is equal to the (instantaneous) forward premia. As a consequence, the instantaneous (nominal) expected return to speculation in the forward foreign exchange market conditioned on available information is zero.

Note that the law of motion for $s(t)$, given by (15) and reproduced again for clarity of exposition:

$$ds(t) = u_2(t)dt + dW_s(t)$$

implies that the conditional expectation of the infinitesimal increment in the spot exchange rate is given by:

$$E_t[ds(t)] = u_2(t)dt$$

therefore, the instantaneous UIP holds if and only if $u_1(t) = u_2(t)$, that is, $u_1(t)$ and $u_2(t)$ are perfectly correlated.

4.2 Conditions for Uncovered Interest Parity in continuous-time models

Let’s go back to the general case where the forward premia $p(t)$ represents the $\tau$-period interest rate differential between, that is, $p(t, \tau)$. Recall that $u(t)$ is defined by the following convolution:

$$u(t) = \int_0^\infty \phi(h)dW_u(t - h)$$

(19)

Recall that the UIP hypothesis implies that the expected change in the spot exchange rate must equal to the forward premia. Therefore, in this continuous time context UIP
is expressed just as:

\[ E_t [\Delta s(t + \tau)] = E_t \left[ \int_0^\tau ds(t + r) \right] = p(t) \]  

(20)

and given that (18) implies that \( u_2(t) \) is the drift of \( s(t) \) we can express the LHS of this last equation as:

\[ E_t \left[ \int_0^\tau ds(t + r) \right] = E_t \left[ \int_0^\tau u_2(t + r)dr + \int_0^\tau dW_s(t) \right] = E_t \left[ \int_0^\tau u_2(t + r)dr \right] \]

Therefore, as in the discrete time case, we need a continuum of forecasts of \( u_2(t + r) \), for the interval \( r \in [0, \tau] \). To this end, we use the Wold decomposition in continuous time. Note that (19) implies:

\[ u_2(t + r) = \int_{-\infty}^{\infty} \phi_{21}(h + r)dW_{u1}(t - h) + \int_{-\infty}^{\infty} \phi_{22}(h + r)dW_{u2}(t - h) \]

To obtain the corresponding expectation conditioned on information available at time \( t \) we need to apply a procedure equivalent to the annihilation operator and zero out \( \phi_{21}(h + r) \) and \( \phi_{22}(h + r) \) for \( t \in [-r, 0] \). In this case, this is straightforward, and we get:

\[ E_t [u_2(t + r)] = \int_{-\infty}^{\infty} \phi_{21}(h + r)dW_{u1}(t - h) + \int_0^\tau \phi_{22}(h + r)dW_{u2}(t - h) \]

Once we have obtained the set of forecasts \( u_2(t) \), we have to aggregate them from 0 to \( \tau \) to obtain the LHS of (20). This yields:

\[ E_t [s(t + \tau) - s(t)] = \int_0^\tau \int_{-\infty}^{\infty} \phi_{21}(h + r)dW_{u1}(t - h)dr + \int_0^\tau \int_{-\infty}^{\infty} \phi_{22}(h + r)dW_{u2}(t - h)dr \]

(21)

On the other hand recall from (19) that \( p(t) \) follows:

\[ p(t) = \int_0^\infty \phi_{11}(h)dW_{u1}(t - h) + \int_0^\infty \phi_{12}(h)dW_{u2}(t - h) \]

Given that both integrals in (21) are defined in wide sense with respect to time we can change the order of integration and equate with this last expression. Then, it is straightforward to see that UIP will hold, that is \( E_t [s(t + \tau) - s(t)] = p(t) \) for every realization of the Wiener processes as long as:

\[ \phi_{11}(h) = \int_0^\tau \phi_{21}(h + r)dr \quad \forall h \]  

(22)

\[ \phi_{12}(h) = \int_0^\tau \phi_{22}(h + r)dr \quad \forall h \]  

(23)
Note that these restrictions do not depend on the sampling period as it was the case of the (discrete-time VAR approach). Therefore, the conditions to be tested are the same whatsoever the election of the sampling period.

5 Example: A continuous-time VAR model

We have derived the conditions that guarantees UIP for a very broad class of models. In this section, we present an example based on the continuous counterpart of a VAR model: the multivariate Orstein-Uhlenbeck. We do so because this continuous time model is a system of linear stochastic differential equations with constant coefficients which, under some regularity conditions\(^5\), generates discrete observations that will satisfy, exactly, a system of stochastic difference equations in which each equation includes the lagged values of all variables in the system, i.e. VAR model. This property will allow us to obtain exact discrete-time representations of the continuous time system which is a useful device for stating UIP conditions as well as for estimation purposes.

Here, the structure is the same as in previous sections, that is:
\[
p(t, \tau) = u_1(t) \\
ds(t) = u_2(t)dt + dW_s(t)
\]
but now \(u(t)\) follows a multivariate Orstein-Uhlenbeck process:
\[
d\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \Phi \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} dt + \begin{pmatrix} dW_{u_1}(t) \\ dW_{u_2}(t) \end{pmatrix} \\
\begin{pmatrix} dW_{1}(t) \\ dW_{2}(t) \end{pmatrix}
\]
where the assumption that \(\Phi\) has two negative eigenvalues guarantees the stationarity of the process.

In particular, note that we can write this model as a continuous time VAR(1) structure:
\[
d\begin{pmatrix} p(t) \\ u_2(t) \\ s(t) \end{pmatrix} = \begin{pmatrix} \Phi & 0 \\ e' & 0 \end{pmatrix} \begin{pmatrix} p(t) \\ u_2(t) \\ s(t) \end{pmatrix} dt + \begin{pmatrix} dW_{u_1}(t) \\ dW_{u_2}(t) \\ dW_{s}(t) \end{pmatrix}
\]
where we can represent it in matrix notation as:

\[ d\xi(t) = A\xi(t)dt + dW(t) \]

and solving this system of linear stochastic differential equations, we get a \( h \)-frequency discrete time representation:

\[ \xi_{t+h} = e^{Ah}\xi_t + \eta_{t+h,h} \]

which is a VAR(1) process in \( \xi_t \), being \( \xi_t \) a discrete time process obtained by sampling \( \{\xi(t), \ t \in [0, \infty)\} \). Moreover, for any matrix \( B \), \( e^B \) is defined by \( e^B = I + \sum_{r=1}^{\infty} \frac{1}{r!} B^r \), and \( \eta_{t,h} = \int_{t-h}^{t} e^{Ah}dW(s) \). This implies

\[
\begin{pmatrix}
  p_{t+h} \\
  u_{2t+h} \\
  s_{t+h}
\end{pmatrix} = \begin{pmatrix}
  e^{Ph} \\
  e^{Ph} \\
  0
\end{pmatrix}
\begin{pmatrix}
  p_t \\
  u_{2t} \\
  s_t
\end{pmatrix} + \begin{pmatrix}
  \eta_{t+h,h} \\
  \eta_{2t+h,h} \\
  \eta_{3t+h,h}
\end{pmatrix} (24)
\]

and rewriting this expression by considering \( \Delta_{h}s_{t+h} \) instead of \( s_{t+h} \)

\[
\begin{pmatrix}
  p_{t+h} \\
  u_{2t+h} \\
  \Delta_{h}s_{t+h}
\end{pmatrix} = \begin{pmatrix}
  e^{Ph} \\
  e^{Ph} \\
  0
\end{pmatrix}
\begin{pmatrix}
  p_t \\
  u_{2t} \\
  \Delta_{h}s_t
\end{pmatrix} + \begin{pmatrix}
  \eta_{t+h,h} \\
  \eta_{2t+h,h} \\
  \eta_{3t+h,h}
\end{pmatrix}
\]

(25)

This expression plays an important role in deriving the conditions that makes UIP hold, as well as for estimation purposes depending on whether \( h = \tau \) (the contract period) or \( h = \delta \) (the sampling period).

### 5.1 Conditions for Uncovered Interest Parity in a continuous-time VAR model

Once you set \( h = \tau \) in (24) we get a first order VAR form for \( x_t = [p_t, u_{2t}, \Delta_{\tau}s_t]^T \)

\[
\begin{pmatrix}
  p_{t+\tau} \\
  u_{2t+\tau} \\
  \Delta_{\tau}s_{t+\tau}
\end{pmatrix} = \begin{pmatrix}
  e^{\Phi_{\tau}} \\
  0 \\
  0
\end{pmatrix}
\begin{pmatrix}
  p_t \\
  u_{2t} \\
  \Delta_{\tau}s_t
\end{pmatrix} + \begin{pmatrix}
  \eta_{t+\tau,\tau} \\
  \eta_{2t+\tau,\tau} \\
  \eta_{3t+\tau,\tau}
\end{pmatrix} (25)
\]

(26)

This discrete-time representation gives the projections of \( x_{t+\tau} \) onto \( x_t \) implied by the continuous-time VAR model. In this case, when the UIP condition holds, then \( E_t [\Delta_{\tau}s_{t+\tau}] = p_t \), and this is only possible in this model if

\[ e^{\Phi_{\tau}} = e^{Ph} \]

so the row corresponding to the projection coefficients of \( \Delta_{\tau}s_{t+\tau} \) onto \( x_t \) is \( (1, 0, 0) \).
5.2 Estimation

Moreover, (24) is also useful from an estimation perspective. Setting $h = \delta$ (the sampling period) and assuming for simplicity, and without loss of generality, that $\delta = 1$ makes

$$
\begin{pmatrix}
  p_{t+1} \\
  u_{2t+1} \\
  \Delta s_{t+1}
\end{pmatrix}
\begin{pmatrix}
  e^\Phi \\
  e^{\Phi-1}(e^\Phi - I) \\
  0
\end{pmatrix}
\begin{pmatrix}
  p_t \\
  u_{2t} \\
  \Delta s_t
\end{pmatrix}
+ 
\begin{pmatrix}
  \eta_{1t+1} \\
  \eta_{2t+1} \\
  \eta_{3t+1}
\end{pmatrix}
$$

(27)

$$\alpha_{t+1} = T\alpha_t + \eta_{t+1}$$

where $\eta_t = \int_{t-1}^t e^{A(t-s)}dW(s)

This system does not allow for direct estimation since $u_{2t+\delta}$ is an unobservable component. However, if we regard (27) as the transition equation of a state space model where $\alpha_t = [u_{1t}, u_{2t}, \Delta s_t]'$ is the state vector and

$$
\begin{pmatrix}
  p_t \\
  \Delta s_t
\end{pmatrix}
\begin{pmatrix}
  1 & 0 & 0 \\
  0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  u_{1t} \\
  u_{2t} \\
  \Delta s_t
\end{pmatrix}
$$

(28)

as the measurement equation we can evaluate the exact Gaussian likelihood via the prediction error decomposition using the Kalman filter as in Harvey and Stock (1989).

To this end, let $Q = E[\eta_t\eta_t'] = \int_0^1 e^{As} \Sigma e^{As'} ds$ while $\alpha_{t|t-1}$ and $P_{t|t-1}$ be the expectation (estimate) of $\alpha_t$ and its covariance matrix conditional on information up to $t - 1$ respectively. Finally, let $\alpha_{t|t}$ and $P_{t|t}$, respectively, be the filtered or predicted values of $\alpha_t$ at time $t$ and its corresponding covariance matrix.

Using this notation, the equations of the prediction step of the basic filter are:

$$\alpha_{t|t-1} = T\alpha_{t-1|t-1}$$

$$P_{t|t-1} = TP_{t-1|t-1}T' + Q$$

$$\eta_{t|t-1} = y_t - y_{t|t-1} = y_t - Z\alpha_{t|t-1}$$

$$f_{t|t-1} = ZP_{t|t-1}Z'$$

while the updating equations are:

$$\alpha_{t|t} = \alpha_{t|t-1} + K_t\eta_{t|t-1}$$

$$P_{t|t} = P_{t|t-1} - K_tZP_{t|t-1}$$

21
Finally, the sample log likelihood can be evaluated via the prediction error decomposition:

\[ \ln L = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln |f_{t|t-1}| - \frac{1}{2} \sum_{t=1}^{T} \eta'_{t|t-1}f_{t|t-1}^{-1}\eta_{t|t-1} \]

where \( N \) is the dimension of \( y_t \), that is, \( N = 2 \).

Moreover, the appendix shows that this state space representation can be expressed as a VARMA(2,1) model and discusses the issue of identification of this model.

6 Final Remarks

Although empirical work have systematically rejected the Uncovered Interest Parity hypothesis, tests on UIP are still an area of active research in international finance, as long as the theoretical studies that have tried to solve the forward premium puzzle have shown a failure at the same level than the rejection of UIP. For this reason it seems important to understand whether this robust empirical fact is related to bad statistical properties of traditional tests on UIP.

In this paper we only examine one dimension of these statistical properties: the temporal aggregation problem. That is, the fact that exchange rates evolve on a much finer time-scale than the frequency of observations. In this set-up where the data sampling frequency is an election of the researcher we care for the impact of this choice on the asymptotic power of tests.

To this end, we have reconsidered tests for UIP in a continuous time framework. In particular we have derived the conditions to test UIP using the Wold decomposition of the continuous-time process and exploiting the (discrete-time) first order companion VAR representation form.

The immediate next step of the paper is conducting a Monte Carlo experiment where data is generated by means of a continuous time model with stochastic volatility, where this last assumption is to capture that exchange rate series often present conditional heteroskedasticity features. Then, through this experiment will be able to calibrate the impact of the choice of a particular sampling frequency into the finite sample size and power of traditional UIP tests. In particular we think that these exercises are going to
be very useful to know which is the maximum degree of overlapping that we can use in
the regression approach without distorting the inference, and to know whether ignoring
the temporal aggregation issue is a problem in the VAR approach.

As future research we plan to extend the model to cover the use of forward exchange
rates with multiple contract dates, \( \tau_1, \tau_2, \ldots, \tau_n \). In this case, the continuous time system
will follow this system of equations:

\[
\begin{align*}
  p(t, \tau_1) &= u_1(t) \\
  p(t, \tau_2) &= u_2(t) \\
  \vdots \\
  p(t, \tau_n) &= u_n(t) \\
  ds(t) &= u_{n+1}(t) + dW_s(t)
\end{align*}
\]

and in particular, we are interested in obtaining the conditions that make:

\[
\begin{align*}
  E_t (s_{t+\tau_1} - s_t) &= p(t, \tau_1) \\
  E_t (s_{t+\tau_2} - s_t) &= p(t, \tau_2) \\
  \vdots \\
  E_t (s_{t+\tau_n} - s_t) &= p(t, \tau_n)
\end{align*}
\]

The advantage of this approach is that we can exploit the time series properties of
the term structure of the forward exchange rate, with the corresponding gains in the
asymptotic efficiency.
Appendix

A Discrete-time representation when \( u(t) \) follows a multivariate Orstein-Uhlenbeck

This appendix presents the derivation of the exact discrete time representation for \( y_t = [p_t, \Delta s_t]' \). In particular, note that (27) implies that:

\[
\begin{align*}
p_t &= f_{11} p_{t-1} + f_{12} u_{2t-1} + \varepsilon_{1t} \\
\Delta s_t &= f_{31} p_{t-1} + f_{32} u_{2t-1} + \varepsilon_{3t}
\end{align*}
\]  

(29)  

(30)

with

\[
u_{2t} = f_{21} p_{t-1} + f_{22} u_{2t-1} + \varepsilon_{2t}
\]

(31)

where \( f_{ij} \) is the element \( i j \) from the matrix \( e^A \).

From (30) we obtain:

\[
u_{2t-2} = \frac{1}{f_{32}} \Delta s_{t-1} - \frac{f_{31}}{f_{32}} p_{t-2} - \frac{1}{f_{32}} \varepsilon_{3t-1}
\]

(32)

Substituting lagged (31) and (32) into (29) and (30), we get obtain a VARMA(2,1) system for \([p_t, \Delta s_t]'\):

\[
\begin{pmatrix}
p_t \\
\Delta s_t
\end{pmatrix} =
\begin{pmatrix}
f_{11} & f_{12} f_{22}/f_{32} \\
f_{31} & f_{22}
\end{pmatrix}
\begin{pmatrix}
p_{t-1} \\
\Delta s_{t-1}
\end{pmatrix} +
\begin{pmatrix}
f_{12}(f_{21} - f_{22} f_{31}/f_{32}) & 0 \\
f_{32} f_{21} - f_{32} f_{31} & 0
\end{pmatrix}
\begin{pmatrix}
p_{t-2} \\
\Delta s_{t-2}
\end{pmatrix} + \eta_t
\]

where

\[
\eta_t =
\begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\varepsilon_{1t} \\
\varepsilon_{2t} \\
\varepsilon_{3t}
\end{pmatrix} +
\begin{pmatrix}
0 & f_{12} & -f_{12} f_{22}/f_{32} \\
0 & f_{32} & -f_{32}
\end{pmatrix}
\begin{pmatrix}
\varepsilon_{1t-1} \\
\varepsilon_{2t-1} \\
\varepsilon_{3t-1}
\end{pmatrix}
\]

Moreover, given this VARMA structure it is easy to show that:

\[
\begin{align*}
\gamma_{21}^\eta(1) &= \frac{f_{32}}{f_{12}} \gamma_{11}^\eta(1) \\
\gamma_{22}^\eta(1) &= \frac{f_{32}}{f_{12}} \gamma_{12}^\eta(1)
\end{align*}
\]

where \( \gamma_{ij}^\eta(1) = E[\eta_{it}\eta_{jt-1}] \), so we can only identify 5 parameters from the variance-covariance matrix \( \Sigma \). In particular, in this paper we assume as an identification scheme that \( dW_a(t) = \alpha_1 dW_{u1}(t) + \alpha_2 dW_{u2}(t) \).
References


