Centre for International Capital Markets

Discussion Papers ***ISSN 1749-3412***

The Implied-Realized Volatility Relation in Foreign Exchange Options Markets

Neil Kellard, Christian Dunis, Nicholas Sarantis

No 2009-19

The Implied-Realized Volatility Relation in Foreign Exchange Options Markets

Neil Kellard[†], Christian Dunis[‡] and Nicholas Sarantis[‡‡]

October 2009

Abstract

Almost all relevant literature has characterized implied volatility as a biased predictor of realized volatility. In this paper we provide new time series techniques to investigate the validity of this finding in several foreign exchange options markets, including the Euro market. First, we develop a new fractional cointegration test that is shown to be robust to both stationary and nonstationary regions. Second, we employ both intra-day and daily data to measure realized volatility in order to assess the relevance of data frequency in resolving the bias. Third, we use data on implied volatility traded on the market. In contrast to previous studies, we show that the frequency of data used for measuring realized volatility within a fractionally cointegrating framework is important for the results of unbiasedness tests. Significantly, for many popular exchange rates, the use of intra-day rather than daily data affects the emergence of a different bias, as the possibility of a fractionally integrated risk premium admits itself!

JEL classification: F31 ; G14 ; C14 ; C22

Keywords: Market efficiency; Options markets; Fractional cointegration; Narrow band least squares; Bootstrap; Traded volatility; Intra-day data

[†]Corresponding author: Essex Finance Centre, Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Tel: 44+1206+87+4153 Email: nkellard@essex.ac.uk

[‡]Centre for International Banking, Economics and Finance (CIBEF), Liverpool Business School, Liverpool John Moores University, John Foster Building, 98, Mount Pleasant, Liverpool, L3 5UZ, United Kingdom. Email: c.dunis@ljmu.ac.uk

[‡‡]Centre for International Capital Markets, London Metropolitan Business School, London Metropolitan University, 84, Moorgate, London, EC2M 6SQ, United Kingdom. E-Mail: n.sarantis@londonmet.ac.uk

1. Introduction

Market efficiency in options markets is typically examined by estimating the following regression

$$\sigma_{t+\tau}^{RV} = \alpha + \beta\sigma_t^{IV} + u_{t+\tau}\,, \tag{1}$$

where σ_t^{IV} is the implied volatility (IV) over a period of time τ and $\sigma_{t+\tau}^{RV}$ is the realized volatility (RV). Unbiasedness holds in (1) when $\alpha = 0$, $\beta = 1$ and $u_{t+\tau}$ is serially uncorrelated. Of course, unbiasedness is a sufficient condition for market efficiency but is not necessary in the presence of either a constant or a time-varying option market risk premium.

Conventional tests in the previous literature have generally led to the conclusion that IV is a biased forecast of RV in the sense that the slope parameter in (1) is not equal to unity (see, *inter alia*, Christensen and Prabhala, 1998, and Poteshman, 2000). This conclusion is found to be robust across a variety of asset markets (see Neely, 2009) and has thus provided the motivation for several attempted explanations of this common finding. Popular suggestions include computing RV with low-frequency data (Poteshman, 2000); that the standard estimation with overlapping observations produces inconsistent parameter estimates (Dunis and Keller, 1995, Christensen et al., 2001); and that volatility risk is not priced (Poteshman, 2000, and Chernov, 2007). However, Neely (2009) evaluates these possible solutions and finds that the bias in IV is not removed.

Of course, the optimality of the estimation procedure applied to (1) depends critically on the order of integration of the component variables. Given the acknowledged persistence in individual volatility series, the recent literature suggests they are well represented as fractionally integrated processes (see, *inter alia*, Andersen, Bollerslev, Diebold and Labys, 2001 and

Andersen, Bollerslev, Diebold and Ebens, 2001). Notably Bandi and Perron (2006), Christensen and Nielsen (2006) and Nielsen (2007) have begun to examine the consequences of this approach for regression (1).

Employing stock market data, Bandi and Perron (2006), Christensen and Nielson (2006) and Nielsen (2007) suggest that IV and RV are fractionally cointegrated series[1]. Interestingly, Bandi and Perron (2006) stress the fractional order of volatility is found in the non-stationary region whereas Christensen and Nielson (2006) and Nielsen (2007) indicate the stationary region. However, allowing for 95% confidence intervals, the estimates could plausibly lie in either region. In any case, Marinucci and Robinson (2001) stress that it is typically difficult to determine the integration order of fractional variables because a smooth transition exists between stationary and non-stationary regions. Christensen and Nielsen (2006) and Nielsen (2007) note that when the fractional nature of the data is accounted for a slope parameter of unity in equation (1) cannot be rejected. Bandi and Perron (2006), noting the non-standard asymptotic distribution of conventional estimators in the non-stationary region, cannot formally test the relevant null hypothesis. However, subsampling shows their results also give support to the unbiasedness hypothesis.

This paper builds on the empirical work of Bandi and Perron (2006), Christensen and Nielsen (2006) and Nielsen (2007) in five steps. Firstly, we employ data for several foreign exchange markets including the relatively new Euro market. Importantly, the IV data collected is traded on the market (and hence is directly observable). Since these data are directly quoted from brokers, they avoid the potential measurement errors associated with the more common approach (see, inter alia, Christensen and Prabhala, 1998) of backing out implied volatilities from a

---

[1]Although this recent work predominantly investigates stock markets, Bandi and Perron (2006) also analyse options on Deutsche Mark/US Dollar futures. Finding similar results to those for stock markets they suggest that fractional

specific option-pricing model.

Secondly, it is important to note that in the recent literature, RV is constructed either from (i) high frequency intra-day return data (see, for example, Nielsen, 2007) or (ii) daily return data (see Bandi and Perron, 2006). Neely (2009) suggests that, at least in the context of least squares regression, the use of intra-day instead of daily data, does not resolve the biased slope coefficient. However, to our knowledge, this comparison has not been formally drawn in a fractionally cointegrated setting. Additionally, given that RV constructed from intra-day data is likely to be a less noisy proxy[2] for the unobserved but true volatility, the key to detecting (small) time-varying risk premia might be the use of such high frequency data. For example, consider augmenting regression (1) with a time-varying risk premium term rp_t

$$\sigma_{t+\tau}^{RV} = \alpha + \beta\sigma_t^{IV} + \delta rp_t + u_{t+\tau} . \qquad (2)$$

Bivariate fractional cointegration between RV and IV implies any risk premium will be of a lower order of (fractional) integration than the original regressors. As a result, and as noted by Bandi and Perron (2006), the use of spectral methods like narrow band least squares will estimate regression (1) consistently, even in the presence of the risk premium. Re-arranging (2) leads to

$$\sigma_{t+\tau}^{RV} - \alpha - \beta\sigma_t^{IV} = \delta rp_t + u_{t+\tau} . \qquad (3)$$

Given that daily data is relatively noisy, it might be that any long memory behaviour of the risk premium[3] is swamped[4] by $u_{t+\tau}$ in finite samples. In other words, a potential pitfall of employing

---

cointegration in the implied-realized relation is a stylised fact.

[2] Anderson and Bollerslev (1998) suggest that daily squared returns are noisy estimators for daily volatility and show that the sum of squared intra-day returns is a less noisy proxy. Employing the theory of quadratic variation, Andersen, Bollerslev, Diebold and Labys (2001) provide theoretical rationale for the intra-day approach.

[3] Kellard and Sarantis (2008) provide evidence for a fractionally integrated risk premium in forward foreign exchange markets. For discussion of volatility risk premia in other markets (see Almeida and Vicente, 2009 and Doran and Ronn, 2008).

daily data to construct RV is that it might render the risk premium undetectable. On the other hand, the use of a less noisy intra-day derived RV may lead to a smaller $u_{t+\tau}$ and therefore the revealing of a time-varying risk premium. Following Bandi and Perron (2006), we deliberately eschew modelling a specific functional form for a risk premium, simply suggesting that fractionally integrated behaviour in the residual of (1) provides prima facie evidence for latent risk premia. To examine these issues, we construct two RV series from intra-day[5] and daily data.

Thirdly, the possibility of fractional cointegration is examined formally using a new adaptation of the recently developed semi-parametric technique of Hassler et al. (2006) [hereafter HMV]. Under certain assumptions HMV prove that a residual-based log periodogram estimator, where the first few harmonic frequencies have been trimmed, has a limiting normality property. In particular, this methodology provides an asymptotically reliable testing procedure for fractional cointegration when the fractional order of regressors present a particular type of non-stationarity. However, given the noted empirical uncertainty, (foreign exchange) volatility may present an integration order that violates the assumptions for the HMV test, as well as other fractional cointegration tests. To circumvent this uncertainty, we suggest, examine and apply an adapted fractional cointegration test robust to both stationary and non-stationary regions.

Fourthly, given the non-standard asymptotic distribution of conventional estimators when using fractionally integrated data, we employ a wild bootstrap procedure as suggested by Gerolimetto (2006) to compute appropriate confidence intervals in (1). Again, this specifically overcomes the difficulties encountered when estimators are applied in the non-stationary region.

Fifthly, we stress that the existence of fractional cointegration and that $\alpha = 0$ and $\beta = 1$

---

[4] For further discussions of swamping in time series see Maynard and Phillips (2001), Kellard (2006) and Kellard and Sarantis (2008).
[5] We thank an anonymous referee for enquiring about the use of intra-day data for constructing RV.

in (1) are only necessary conditions for unbiasedness. The important condition, that $u_{t+\tau}$ in (1) is serially uncorrelated, is required but such tests have been neglected by the recent extant literature. For completeness therefore, we employ an appropriate portmanteau test to the fractionally cointegrating residual.

The paper is divided into five sections: Section 2 presents the empirical methodology; section 3 describes the data; section 4 analyses the empirical results and, finally, section 5 concludes.

2. Empirical methodology

2.1. Fractional integration

Many in the literature (see, *inter alia*, Bandi and Perron, 2006, Vilasuso, 2002, Andersen, Bollerslev, Diebold and Labys, 2001 and Baillie et al. 1996) have suggested that asset price volatility is neither an I(1) nor an I(0) process but rather a fractionally integrated or I(*d*) process. The introduction of the autoregressive fractionally integrated moving average (ARFIMA) model by Granger and Joyeux (1980) and Hosking (1981) allows the modelling of persistence or long memory where $0 < d < 1$. A time series y_t follows an ARFIMA[6] (*p, d, q*) process if

$$\Phi(L)(1-L)^d y_t = \mu + \Theta(L)\varepsilon_t, \quad \varepsilon_t \sim iid(0,\sigma^2); \tag{4}$$

where $\Phi(L) = 1-\phi_1 L - ... - \phi_p L^p$ and $\Theta(L) = 1-\theta_1 L - ... - \theta_q L^q$. Such models may be better able to describe the long-run behaviour of certain variables. For example, when $0 < d < 1/2$, y_t is

[6] ARFIMA models have been often used to model and forecast volatility (see, *inter alios*, Konstantinidi et al., 2008 and Becker et al., 2007).

stationary but contains long memory, possessing shocks that disappear hyperbolically not geometrically. Contrastingly, for $1/2 < d < 1$, the relevant series is non-stationary, the unconditional variance growing at a more gradual rate than when $d = 1$, but mean reverting.

The memory parameter d can be estimated by a number of different techniques. The most popular, due to its semi-parametric nature, is the log periodogram estimator (Geweke and Porter-Hudak, 1983; Robinson, 1995a) henceforth known as the GPH statistic[7]. This involves the least squares regression

$$\log I(\lambda_j) = \beta_0 - d \log\{4\sin^2(\lambda_j / 2)\} + u_j, \quad j = l+1,\ l+2, ..., m; \tag{5}$$

where $I(\lambda_j)$ is the sample spectral density of y_t evaluated at the $\lambda_j = 2\pi j / T$ frequencies, T is the number of observations and m is small compared to T. Inter alia, Robinson (1995), Pynnönen and Knif (1998) and HMV, note that the least-squares estimate of d can be used in conjunction with standard t-statistics. For the stationary range, $-1/2 < d < 1/2$, Robinson (1995) demonstrated that the GPH estimate is consistent and asymptotically normally distributed. Additionally, Velasco (1999) shows that when the data are differenced, the estimator is consistent for $1/2 < d < 2$ and asymptotically normally distributed for $1/2 < d < 7/4$.

2.2. Fractional cointegration

As discussed in the introduction, some recent literature has presented the possibility that RV and IV are fractionally cointegrated. Fractional cointegration can be defined by supposing y_t

[7] It should be noted that some semi-parametric estimators, for example the Gaussian semi-parametric (GSP) estimator, can be more efficient than the GPH alternative. However, given the popularity of the GPH estimator and its use by HMV, we employ the log periodogram procedure. In any case, recent work has suggested that alternative semiparametric estimators are only more efficient than the GPH estimator under certain conditions (see Baillie and Kapetanios, 2008).

and x_t are both I(d), where d is not necessarily an integer and the residuals, $u_t = y_t - \beta x_t$, are I($\delta = d - b$). When $0 < b < d$, where b is also not necessarily an integer, series are fractionally cointegrated. Testing for fractional cointegration can be accomplished using a multi-step methodology (see HMV) where (i) the order of integration of the constituent series are estimated and tested for equality and (ii) the long-run equilibrium relationship is estimated[8] and the residuals examined for long-memory. Alternative methodologies include the joint estimation of memory parameters of the constituent series, the cointegrating residuals and the equilibrium relationship (see Velasco, 2003 and Nielsen, 2007) or the use of bootstrap methods (see Davidson, 2002 and 2005).

A frequently used approach is to adopt a multi-step methodology where the concluding step estimates the GPH statistic, δ , for the least squares residual of the equilibrium relationship (see Dittmann, 2001). Inter alia, Tse et al. (1999) experimentally noted that t-statistics associated with $\hat{\delta}$ might not be normally distributed.

2.3. Nonstationary fractional cointegration

If d is in the *strongly* non-stationary region, HMV demonstrate that as long as $\hat{\beta}$ can converge fast enough, $\hat{\delta}$ possesses a limiting normal distribution provided the very first harmonic frequencies are trimmed. Specifically, this entails setting $l > 0$ in (5). Moreover, Monte Carlo experiments show that trimming only one frequency, $l = 1$, provides a satisfactory

[8]The long-run equilibrium relationship itself could be approximated by OLS, a fractional version of the Fully Modified method suggested by Kim and Phillips (2001), Gaussian semi-parametric estimation developed by Velasco (2003) or narrow band spectral estimates (see Robinson and Marinucci, 2003).

normal approximation for the distribution of GPH statistic in finite samples. Of course, given (asymptotically) normal estimators, standard inference procedures can be legitimately applied.

A priori, it is useful to note the HMV test theoretically requires certain assumptions to hold to generate limiting normality for the distribution of $\hat{\delta}$. The most relevant to our discussion are listed below

$$0 \le \delta < 0.5, \tag{6}$$

$$\delta < d - 0.5 < 1, \tag{7}$$

$$0.72 < d < 1.5. \tag{8}$$

In particular, it should be stressed that condition (8) implies that for what might be termed the *weakly* non-stationary region (i.e. $0.5 < d < 0.72$), there is no limiting normal distribution theory. This is due to the slower convergence rate of $\hat{\beta}$.

2.4. Stationary fractional cointegration

If $d < 0.5$, OLS estimates of β are inconsistent suggesting the above approach may be inappropriate. However, Robinson and Marinucci (2003) and Christensen and Nielsen (2006) have shown that narrow band least squares (NBLS) estimation can result in an estimator $\hat{\beta}_z$ that is consistent and normally distributed. To explain NBLS consider first that a matrix form of (1) could be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{9}$$

where $\boldsymbol{\beta}$ is a 2×1 vector of unknown coefficients and $\mathbf{u}$ is a $T \times 1$ vector of disturbances. Additionally define the complex $T \times T$ Fourier matrix, $\mathbf{V}$, which has as its *(j,t)th* element

$$v_{jt} = T^{-1/2}\left[\exp\{-i\lambda_{j-1}t\}\right], \quad j, t = 1,...,T \tag{10}$$

and presents the frequencies

$$\lambda_j = \frac{2\pi j}{T}. \tag{11}$$

A transformation to the frequency domain (see Harvey, 1993) can be made by pre-multiplying the observation matrices in (9) by $\mathbf{V}$ and expressing the transformed model as

$$\dot{\mathbf{y}} = \dot{\mathbf{X}}\boldsymbol{\beta} + \dot{\mathbf{u}}, \tag{12}$$

where $\dot{\mathbf{y}} = \mathbf{Vy}$, $\dot{\mathbf{X}} = \mathbf{VX}$ and $\dot{\mathbf{u}} = \mathbf{Vu}$. OLS estimation of (12) will produce identical estimates to that of (1). Note however different frequency components may be omitted by removing $T - z$ corresponding transformed observations. This is band spectrum regression and it can be shown that the β in (1) will be estimated by the statistic

$$\hat{\beta}_z = \left[\sum_{j=0}^{z} I_{\sigma_{IV}}(\lambda_j)\right]^{-1} \sum_{j=0}^{z} I_{\sigma_{IV}\sigma_{RV}}(\lambda_j), \quad 0 \le z \le T-1; \tag{13}$$

where $I_{\sigma_{IV}}(\lambda_j)$ is the sample spectral density of IV and $I_{\sigma_{IV}\sigma_{RV}}(\lambda_j)$ is the cross-spectrum between IV and RV[9]. Additionally, such band spectrum regression is called NBLS if

$$\frac{1}{z} + \frac{z}{T} \to 0 \;\; as\, T \to \infty. \tag{14}$$

A number of assumptions are required for the generation of a limiting normal distribution for $\hat{\beta}_z$. These include

$$d > 0, \tag{15}$$

$$\delta \ge 0, \tag{16}$$

$$d + \delta < 0.5. \tag{17}$$

---

[9]Therefore, $\hat{\beta}_{T-1}$ is a special case, equivalent to the OLS estimate of β in (1).

Christensen and Nielsen (2006) recently employed a multi-step methodology, where the concluding step semi-parametrically estimates δ for the NBLS residual of the equilibrium relationship. Hypothesis testing is then conducted on $\hat{\delta}_z$ as if the residuals were observed. Although possibly an appropriate procedure (see Nielsen, 2007) this has not been ascertained in the literature[10].

2.5. Boundary fractional cointegration

The fractional order of volatility has typically been found to have confidence intervals that span the stationary/non-stationary boundary (i.e. $0.3 < d < 0.7$). In any case, Marinucci and Robinson (2001) suggest determining the stationarity or otherwise of time series variables is often difficult. This difficulty is particularly pronounced in a fractional context, where a smooth transition exists between stationary and non-stationary regions. However, the use of the fractional cointegration methodology discussed above relies on the identification of the appropriate region. Furthermore, point estimates for d are often found in the weakly non-stationary region (i.e. $0.5 < d < 0.72$), a result for which there is no limiting normal distribution theory. Therefore, we require an estimator which is robust in finite samples to orders of integration for d that span the boundary. As a first step, it would seem more appropriate to use NBLS rather than OLS to estimate the equilibrium relationship. As discussed above NBLS, in contrast to OLS, provides a consistent estimator in the stationary region. However, analogously

[10]However, Nielsen (2007) uses a local Whittle QMLE to jointly estimate the integration order of the regressors, the integration order of the residuals and the coefficients of the cointegrating vector. Under certain conditions, this estimator is shown to be asymptotically normal for the stationary region.

to OLS, NBLS is consistent in the non-stationary region (see Marinucci and Robinson, 2001). Furthermore, in the non-stationary region, NBLS generally converges faster than OLS (see Marinucci and Robinson, 2001) and thus may resolve the lack of a normal distribution for $\hat{\delta}$ in the weakly non-stationary region shown by HMV. In a second step, it would appear useful to examine the effect of trimming on the distribution of $\hat{\delta}_z$.

To assess the distribution of $\hat{\delta}_z$ across the boundary we perform a simple simulation. Let x_t be generated by an ARFIMA $(0, d, 0)$ series

$$(1-L)^d x_t = \varepsilon_{1t}, \tag{18}$$

where the fractional difference operator is defined by the Maclaurin series

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(-d+j)L^j}{\Gamma(-d)\Gamma(j+1)} = \sum_{j=0}^{\infty} d_j L^j, \;\; d_j = \frac{(j-1-d)d_{j-1}}{j}, \;\; d_0 = 1 \tag{19}$$

and $\Gamma(.)$ is the gamma function. To avoid the initial conditions effect, sample sizes $t = 1,...,T+w$ are generated and the first $w = 1000$ observations removed. Additionally, $\sum_{j=0}^{\infty} d_j L^j$ is approximated by allowing $d_j = 0$ when $j > 1000$. The true regression model is

$$y_t = x_t + v_t, \;\; (1-L)^{\delta} v_t = \varepsilon_{2t} \tag{20}$$

and is estimated by NBLS

$$y_t = \hat{\alpha}_z + \hat{\beta}_z x_t + \hat{v}_t. \tag{21}$$

GPH statistics $\hat{\delta}_z$ are then computed for $\hat{v}_t$ and the statistics below calculated

$$\frac{\hat{\delta}_z - \delta}{se(\hat{\delta}_z)}. \tag{22}$$

In the experiments two-sided tests at the 1%, 5% and 10% level are calculated. To allow comparison with HMV we set $m = T^{0.5}$, allowed the trimming parameter $l \in (0,1)$ and use 2000 replications. Indeed, simulations not reported here, and using OLS instead of NBLS, replicate the Monte Carlo results of HMV. In our reported simulations we also vary the NBLS estimation by applying both $z = T^{0.3}$ and $z = T^{0.75}$. Recent work has recommended the use of a low number of frequencies (see Christensen and Nielsen, 2006, Marinucci and Robinson, 2001, and Robinson and Marinucci, 2003) and thus the two settings will allow some assessment of this approach. To begin with Tables 1 and 2 show the size of the GPH tests, without trimming, employing different values of d and δ.

[Insert Tables 1 and 2]

The results above clearly show that NBLS/GPH methodology, without trimming, does not typically produce a normally distributed test statistic. Tests are particularly oversized when a relatively small number of frequencies ($z = T^{0.3}$) are employed[11]. This casts some doubt on the recent approach used in the literature and the twin assertions that z should be relatively low and hypothesis testing can be conducted on $\hat{\delta}_z$ as if the residuals were observed. Tables 3 and 4 below repeat the previous simulations, although now with $l = 1$.

[Insert Tables 3 and 4]

Tables 3 and 4 clearly show that with trimming, the new methodology produces an approximately normally distributed test statistic. Strikingly, this result holds for when d is in the stationary region or the weakly non-stationary region. The reduction in size bias is particularly marked in the $z = T^{0.3}$ case. These finite sample results have clear implications for testing option market efficiency using regression (1). They suggest that as long as trimming is employed an

[11]Marinucci and Robinson (2001) also examine the NBLS estimate of the cointegrating vector and associated semiparametric methods for testing for the existence of fractional cointegration. Using a Monte Carlo approach, as one would expect, their Hausman test is shown to be similarly oversized. It should be noticed that their Monte Carlo investigation only examined the 5% size over the strongly non-stationary region (i.e. $0.8 < d < 1.2$). See Marinucci and Robinson (2001) Table 11.

NBLS/GPH methodology can be legitimately used to assess whether RV and IV are fractionally cointegrated. We shall use this result in empirical tests of data discussed below.

3. Data

Appropriate time series of RV and IV were constructed for the period January 1991[12] to December 2007. As in Dunis and Keller (1995), Dunis and Huang (2002), and Sarantis (2006), IV is measured by at-the-money, one-month forward, market quoted daily volatilities at close of business in London, obtained from brokers by Reuters[13]. These 'traded' implied volatilities[14] measure the market's expectation about the future volatility of the spot exchange rate for six currencies: Sterling/US dollar, US dollar/Swiss Franc, US dollar/Yen, Euro/Yen, Euro/Sterling and Euro/US dollar[15]. As currency volatility has now become a traded quantity in financial markets, it is therefore directly observable on the marketplace. The databank is maintained by CIBEF at Liverpool John Moores University. As noted in the introduction, since these data are

---

[12]The choice of start date was governed by the availability of IV data.

[13] Traded volatilities were originally exclusively at-the-money volatilities: option traders would then use at-the-money volatilities to price different strikes (including out-of the money strikes), and out-of-the money volatilities were just backed out from option pricing models where the price, strike, maturity and interest rate are known and the model solved for the volatility (i.e. this is how the volatility smile would be obtained). Combined with the fact that at-the-money volatilities are by far the most liquid and are what volatility traders are interested in keeping, Reuters started recording only at-the-money volatilities since the early 1990s. More recently, for very liquid assets like some FX rates, Reuters has started recording several out-of-the money volatilities (with 8 delta points for both puts and calls). However there is only 2 years history available and it is very difficult to extract the data in a consistent time series format suitable for econometric estimation.

[14]Implied volatilities are also annualised rates so that a quoted volatility of 5 per cent typically translates to a monthly variance rate of $(0.05^2)(21/252)$. The calculations assume that annualised rates refer to a 252 trading day year.

[15]The sample period for the three Euro exchange rates begins in January 1999, when the Euro was introduced. See Yang et al. (2008) for further discussion of Euro exchange rates.

directly quoted from brokers, they avoid the potential biases associated with backing out.

Given IV data for each specific day, two measures of RV are calculated over the remaining one month of the option. In the extant literature, RV has been recently derived from either intra-day return data or daily returns data. Firstly, using high-frequency data (sourced from Olsen Associates), a 5-minute logarithmic return series is calculated for each foreign exchange rate series[16]. From the sum of the 5-minute squared returns, the daily variance (h_t) is computed and then the RV quantity

$$\sigma_{t+\tau}^{RV^h} = \sqrt{\frac{252}{\tau - 1}\sum_{i=1}^{\tau} h_{t+i}} \ . \qquad (23)$$

Secondly, following Christensen and Hansen (2002) and using daily returns data

$$\sigma_{t+\tau}^{RV^d} = \sqrt{\frac{252}{\tau - 1}\sum_{i=1}^{\tau} (r_{t+i} - \bar{r}_t)^2} \ , \qquad (24)$$

where $r_t = \ln(S_t / S_{t-1})$, τ is the relevant number of trading days[17] and S_t is the closing (London time) average of bid and ask quotes for the spot exchange rates. The raw daily dataset thus consists of (3338) 4348 time series observations for each (euro) volatility series. Of course, as pointed out by Christensen and Prabhala (1998), overlapping data problems will beset estimation of equation (1) if daily datasets are employed. To circumvent this a monthly dataset is derived from the daily version. Specifically, IV data is taken only from the subsequent trading day after the final day used in the calculation of the previous RV figure. This allows the data to cycle through the calendar and the resulting dataset contains (106) 198 non-overlapping observations for each (euro) volatility series[18].

[16] For a discussion on foreign exchange market efficiency when employing tick frequency data see Akram et al. 2009.

[17] Assumed to be 21 days.

[18] Similarly in Bandi and Perron (2006) the monthly dataset contains 152 observations. However, IV data is taken

4. Empirical results

GPH statistics[19] for the logarithm[20] of monthly[21] volatility series were estimated using differenced data[22] and Ox version 3.3 (see Doornik, 1999) and are shown in Table 5. An alternative approach would be to specify fully parametric ARFIMA (p,d,q) models computed by exact maximum likelihood (EML). Of course, this fully parametric approach is more efficient but will be inconsistent if the short-run dynamics are incorrectly specified[23].

[Insert Table 5]

---

only from the closing value of each month. Although common practice, particularly in forward market analysis, this methodology does not ensure that periods of observation are strictly non-overlapping. For example, an IV figure drawn from the last trading day in January 1991 (Thursday 31st) would be matched with a RV figure calculated from 21 days of subsequent trading day returns (i.e. data up to and including Friday 1st March). Of course the next IV figure would be drawn from the last trading in February (Thursday 28th) causing subsequent periods of observation to overlap. In contrast, the cycling dataset suggested here ensures the non-overlapping nature of the data in construction. Additionally, the cycling dataset does not draw data solely from one period of the month and therefore is not likely to be as susceptible to any intra-monthly seasonality. See Breuer and Wohar (1996) for an analogous application of cycling monthly datasets to the forward foreign exchange market.

[19]Note that the GPH statistic was estimated at $m = T^{0.75}$ following Maynard and Phillips (2001) and Kellard and Sarantis (2008). To maintain consistency with the previous section we also estimated the GPH statistic using $m = T^{0.5}$ which produced similar point estimates but larger standard errors. As we have data at a monthly frequency, the use of smaller bandwidths produces standard errors that are notably less informative over the orders of integration we are interested in. Given the apparent lack of bias at the larger bandwidth we decided to work with $m = T^{0.75}$. Moreover, the estimated standard error of d is that derived by Geweke and Porter-Hudak (1983) and shown in equation (7) of HMV, who show it to be more appropriate than the conventional and Robinson (1995) alternatives.

[20]Natural logarithms of all volatility series were taken to minimise the possibility of non-normal variables as shown by, inter alia, Christensen and Hansen (2002).

[21]All empirical analysis is carried out on the monthly dataset to avoid the overlapping data problems discussed by Christensen and Prabhala (1998).

[22]The resulting estimate of d was then increased by 1. If $\hat{d} < 0.5$ then d was re-estimated using data in levels. Also note that in (5), l is set equal to zero, indicating no trimming of the harmonic frequencies.

[23]Recent work by Nielsen (2007), Christensen and Nielsen (2006) and Bandi and Perron (2006) all employ semi-parametric estimation of the long memory parameter.

Table 5 contains some interesting results. Firstly, the GPH point estimates of fractional differencing in foreign exchange volatility are spread over the range 0.87 to 0.30. Tests for $d = 1$ and $d = 0$ show that, in particular, the volatility series are typically[24] fractionally integrated with $0 < d < 1$. These results confirm that foreign exchange market behaviour is analogous to stock market behaviour investigated by Bandi and Perron (2006), Christensen and Nielsen (2006) and Nielsen (2007). Secondly, standard errors are such that it cannot be ascertained whether volatility series are characteristically stationary or non-stationary fractionally integrated processes. Thirdly, the point estimates for RV calculated from high frequency data (RV^h) are, for 5 of the 6 currencies, closer to those of IV than point estimates for RV derived from daily data (RV^d). This gives an early indication that the choice of data frequency when constructing RV may have significant consequences within a long memory framework; in particular, it shows that it may be preferable to employ intra-day rather than daily data. Fourthly, RV and IV series appear to have similar orders of integration. To examine this in more detail we test that the fractional orders of the constituent variables are equal by applying the homogenous restriction

$$H_0 \ : \ PD = 0 , \tag{25}$$

where $D = \begin{bmatrix} d_{RV} \\ d_{IV} \end{bmatrix}$ and $P = \begin{bmatrix} 1 & -1 \end{bmatrix}$. Robinson (1995) noted the relevant Wald test statistic could be expressed as

$$\hat{D}' P' \left[(0, P) \left\{ (Z'Z)^{-1} \otimes \Omega \right\} (0, P)' \right]^{-1} P \hat{D} , \tag{26}$$

where Ω is residual variance-covariance matrix from (5), $Z = [Z_{l+1} ... Z_m]'$ and $Z_j = [1, -\log\{4 \sin^2(\lambda^2 / 2)\}]$. Table 6 contains the Wald test results.

[24] Of the 18 volatility series, all rejected the null $d = 0$ and only three could not reject $d = 1$.

[Insert Table 6]

Employing six currencies with two RV series each, there are 12 RV-IV pairs. For only one pair (RV^d, IV) US$/Yen are equal fractional orders of *d* rejected (with a p-value of 0.03). However, for the same currency, (RV^h, IV) presents a p-value of 0.28 and therefore the Wald test does not reject the null. Given the strong evidence for equal fractional orders for RV and IV in a foreign exchange context, for completeness we continue to use all the RV-IV pairs in the subsequent analysis.

It is now useful to examine the fractional differencing parameter of the possible cointegrating relationship. Of course, given that volatility series clearly have confidence intervals for d that typically span the stationary/non-stationary boundary it would appear sensible to employ a fractional cointegration test robust to both these regions. Thus we next apply the adapted test proposed and examined in section 2.5. Specifically, in a first step, regression (1) is estimated by NBLS and employing $z = T^{0.75}$. In a second step, the NBLS residuals from (1) are tested for their order of integration using GPH with $l = 1$. The resulting estimates of δ, are shown below in Table 7.

[Insert Table 7]

Interestingly, the point estimate of δ is always lower than the fractional parameter d of the constituent series, implying bi-variate fractional cointegration. However, for 3 out of 6 (RV^h, IV) pairs the null $\delta = 0$ *can* be rejected (i.e. UK£/US$, US$/SF and Euro/Yen). The point estimate of δ in these rejection cases ranges from 0.254 to 0.411 and suggests the possibility of a (small) time-varying risk premium[25], exhibiting long memory behaviour in the stationary region. This result provides evidence against unbiasedness in a fractionally integrated framework. Similar evidence has recently been found for a fractionally integrated risk premium in the forward foreign exchange market (see Kellard and Sarantis, 2008). By contrast, for 5 out of 6 (RV^d, IV) pairs the null $\delta = 0$ *cannot* be rejected[26]; the additional noise when RV is constructed from daily data, masking the long memory evidenced with higher frequency data.

In any case, for currency pairs above presenting fractional cointegration with $\delta = 0$, this is only a necessary but not sufficient condition for unbiasedness in the options foreign exchange market. Therefore we need to consider the intercept and slope parameters in (1). As a preliminary step and for comparative purposes we present conventional OLS estimates in Table 8.

[Insert Table 8]

These suggest that, as has previous literature, IV is generally a biased predictor of RV in the foreign exchange market. However, the slope coefficients found are generally much closer to unity than those estimated in previous studies (see Neely, 2009). As we are employing traded volatility for the first time, this suggests that perhaps the measurement error in 'backing out' implied volatility from option pricing models may have more effect on biasing parameters than previously acknowledged. Furthermore, the use of high frequency data to construct RV again makes a difference to the interpretation of results. Strikingly, the US dollar exchange rates present more downward bias when using (RV^h, IV) than (RV^d, IV) pairs. Conversely, the Euro exchange rates reveal less bias when employing (RV^h, IV) pairs, with slope coefficients of around unity. This is pertinent, because in an exchange rate context at least, empirical evidence

[25] The question now arises, why would some currencies present a fractionally integrated risk premium and others not? One possibility, given the unobservable nature of true volatility, is that in certain cases even proxy RV derived from intra-day data is still too noisy to reveal the small latent premia.

[26] For the Euro/UK£ rate, the (RV^d, IV) pair show a $\hat{\delta} = -0.41$. Such anti-persistence is not commonly found in financial time series and may be due to the data limitations inherent in employing the relatively short Euro series. Moreover, the 'superior' (RV^h, IV) pair gives an estimate of δ which is not significantly different from zero.

for the well known biased slope coefficient is often derived from US dollar rates. It would appear this *form of bias* may not be as severe in a Euro context.

Of course, as already noted earlier, the fractional order of integration of volatility is likely to have an effect on the OLS estimation of (1). In particular, if $d < 0.5$ then OLS estimates will be inconsistent. However, even if $d > 0.5$, OLS will typically converge slower than NBLS. Therefore, Table 9 provides the NBLS[27] point estimates for (1).

[Insert Table 9]

NBLS parameter estimates are consistent but have non-standard limit distributions in the non-stationary region. To circumvent this, and following Gerolimetto (2006), a wild bootstrap procedure[28] is employed to generate 90% and 95% confidence intervals for the slope coefficient in (1). Specifically, in the frequency domain, NBLS residuals $\dot{\mathbf{u}}$ are resampled with replacement and used to generate a bootstrapped dependent variable $\dot{\mathbf{y}}^*$. The new dependent variable is regressed on the original frequency domain regressors $\dot{\mathbf{X}}$ to get the bootstrapped coefficient vector $\boldsymbol{\beta}^*$. Using the bootstrap class in OX, 1000 bootstrapped slope coefficients were generated by this procedure.

Strikingly, Table 9 shows that the use of (RV^h, IV) provides strong, if not ubiquitous, evidence for unity slope coefficient. However, it should be noted that for the 2 of the 6 currencies, the Euro/Yen[29] and the UK£/US$, a unity coefficient is outside the 95% and 90%

[27] Again employing $z = T^{0.75}$.

[28] Gerolimetto (2006) notes that alternatives to the wild bootstrap, for example the moving block bootstrap and subsampling, present the disadvantage that performance can depend on the ad hoc selection of the number of blocks or subsamples.

[29] The estimated greater than unity slope coefficient for the Euro/Yen (RV^h, IV) pair is a much less frequent empirical finding than the common downward bias. However, recent literature has reported similar coefficients (see Chernov, 2007). Again, this may be due to the data limitations inherent in employing the relatively short Euro series.

confidence intervals respectively. Clearly, for all (RV^d, IV) pairs, the NBLS slope parameter is much closer to unity than the OLS version. In a similar vein, the NBLS constant approaches zero. Furthermore, the 90% and 95% confidence intervals for the NBLS slope coefficient all include unity.

It is imperative to note that fractional cointegration with $\alpha = 0$ and $\beta = 1$ in (1) are still only necessary, not sufficient conditions for unbiasedness. It is also required that $u_{t+\tau}$ in (1), the forecast error, is serially uncorrelated but tests of this condition are typically absent in the recent extant literature. As noted earlier, the finding of a fractionally integrated cointegrating residual for 3 (RV^h, IV) pairs (i.e. UK£/US$, US$/SF and Euro/Yen), suggests a time-varying risk premium and provides evidence against unbiasedness in these cases[30]. But does unbiasedness hold for other currency pairs? Defining $u_{t+\tau} = \sigma_{t+\tau}^{RV} - \sigma_t^{IV}$ we assess dependence in $u_{t+\tau}$ by employing a test based on the set of sample autocorrelations

$$r_j = \frac{\sum_{t=j+1}^{T} \left(u_{t+\tau} - \bar{u}\right)\left(u_{t+\tau-j} - \bar{u}\right)}{\sum_{t=1}^{T} \left(u_{t+\tau} - \bar{u}\right)^2}, \qquad j = 1, 2, ...; \qquad (27)$$

where T is the number of observations. If $u_{t+\tau}$ is homoscedastic and white noise, it is well known that r_j is asymptotically normally distributed with mean 0 and variance

$$v_{1j} = \left[T(T+2)\right]^{-1}(T - j), \quad j = 1, 2, ... \qquad (28)$$

[30] Again, for completeness we continue to use all the RV-IV pairs in the subsequent analysis.

However, *inter alios*, Lo and MacKinlay[31] (1988, 1989) and Islam and Khaled (2005) note that a heteroscedasticity-consistent estimator of the variance of the sample autocorrelations r_j is instead provided by

$$v_{2j} = \frac{\sum_{t=j+1}^{T} \left(u_{t+\tau} - \bar{u}\right)^2 \left(u_{t+\tau-j} - \bar{u}\right)^2}{\left[\sum_{t=1}^{T} \left(u_{t+\tau} - \bar{u}\right)^2\right]^2}, \qquad j = 1, 2,..... \qquad (29)$$

Of course, even if $u_{t+\tau}$ were white noise, it would be unsurprising to find statistical significance in at least one of the first twenty sample autocorrelations. For this reason, it is general practice to evaluate the first m sample autocorrelations as a set. One approach is through the portmanteau test of Ljung and Box (1978), which examines whether, taken as a set, $r_j \left(j = 1,\ldots,m \right)$ are too large in absolute value to support the hypothesis that $u_{t+\tau}$ is white noise. The test statistic is therefore

$$Q_m = \sum_{j=1}^{m} v_{2j}^{-1} r_j^2, \qquad m = 1,2,..., M; \qquad (30)$$

where under the random walk null hypothesis, Q_m has an asymptotic chi-square distribution with m degrees of freedom. The relevant p-values for the test can be found in Tables 10a and 10b.

[Insert Table 10]

Table 10a provides results for the (RV^h, IV) pairs. In 4 of the 6 currencies, p-values of zero provide strong evidence for serial correlation in $u_{t+\tau}$. Notably, 3 of these currencies also presented a fractionally integrated cointegrating residual and the portmanteau test underscores this rejection of unbiasedness. However, for two currencies (Euro/US$ and to a lesser extent,

[31] See, for example, Lo and MacKinlay (1988) equation (19). Note that the same estimator of the variance of the sample autocorrelations is also employed by Lo and MacKinlay, 1989, Newbold et al., 1998 and Lobato et al.,

US$/Yen) a large majority of p-values indicate insignificance at the 10% level. These two currencies therefore present modest evidence for unbiasedness. Turning to Table 10b, results for (RV^d, IV) pairs again provides a clear contrast with prior findings. Now for 5 of the 6 currencies virtually all of the p-values indicate insignificance at the 10% level, suggesting that $u_{t+\tau}$ is serially uncorrelated. The exception is the UK£/US$ rate, where the majority of p-values indicate significance at the 10% level. Overall, it would appear that the use of intra-day data reveals structure in the forecast error, whereas daily data veils this structure. In an implied-realized volatility context, we therefore posit it is clearly preferable to employ a high frequency construction of RV.

5. Conclusions

Almost all relevant literature has characterized implied volatility (IV) in the foreign exchange options markets as a biased predictor of realized volatility (RV). The cause of this bias has been the subject of much debate but in a recent paper, Neely (2009), the popular suggestions (i.e. overlapping data; use of low frequency data; and the non-pricing of volatility premia) are rejected.

In this paper we examine the unbiasedness hypothesis by employing data for several very liquid foreign exchange options markets, including the relatively new Euro market. We make three contributions to the existing literature. Firstly, we develop, examine and apply a new test for fractional cointegration which is shown to be robust to both the stationary and nonstationary regions, including the weakly non-stationary region. Secondly, given that RV constructed from

---

2001.

intra-day data is a less noisy proxy for the unobserved but true volatility, we posit that any (small) time-varying risk premia are more likely to be detected using the higher frequency data. To examine this issue, for each currency we construct two RV series from intra-day and daily data respectively and their respective relation with IV is compared throughout the empirical analysis. Thirdly, we collect data on IV that is traded on the market (and hence is directly observable). Since these data are directly quoted from brokers, they avoid the potential measurement errors associated with the more common approach of backing out implied volatilities from a specific option-pricing model.

In contrast to previous studies, we find that the frequency of data used for the construction of RV is important for both the results of unbiasedness tests and the detection of time-varying risk premia in foreign exchange options markets. Employing the new fractional cointegration test, we show that foreign exchange RV and IV are fractionally cointegrated across a range of currencies and data frequencies. Moreover, tests using bootstrapped estimates and confidence intervals are not typically able to reject the hypothesis that the slope parameter in the RV-IV relation is unity. Contrary to the widely held view derived from previous research, the slope coefficient in the RV-IV relation is therefore shown not to be the primary source of bias.

However, serial correlation tests of the forecast error in the RV-IV relation, which are frequently neglected in the extant literature, reveal a different picture. Results based on the RV measure derived from intra-day data show significant serial correlation in the forecast error for four out of six currencies, which suggests rejection of the unbiasedness hypothesis. In contrast, results obtained from the daily RV typically indicate absence of serial correlation in the forecast error and hence they support unbiasedness. It would appear that intra-day data reveal structures in the forecast error, while the more noisy and hence less reliable daily data veils this structure.

Furthermore, when we employ the RV measure derived from intra-day data, the forecast errors for currencies like the Sterling/US$, Euro/Yen and US$/Swiss Franc, are shown to be fractionally integrated, suggesting the possibility of a time-varying risk premium with long memory. This is a new finding in the literature on options markets. This result is not found when the RV measure is derived from daily data and provides further testimony that the use of less noisy proxies for RV has significant time series implications.

Acknowledgements

Previous versions of this paper were presented at the European Financial Management Association 2007 Annual Conference (Vienna University of Economics and Business Administration), the International Workshop on Computational and Financial Econometrics 2007 (University of Geneva) and a seminar at the University of Liverpool Management School (2006). Helpful comments from the participants at these presentations are acknowledged. We would also like to thank Jerry Coakley, Ana-Maria Fuertes, David Harvey, Petko Kalev, Brendan McCabe, John Nankervis, Anne Peguin-Feissolle and Mark Wohar for their encouragement, comments and support. We are particularly grateful to an anonymous referee for their constructive comments and suggestions. The opinions expressed in the paper and any remaining errors remain our responsibility.

References

Akram, Q., Rime, D., Sarno, L., 2009. Does the law of one price hold in international financial markets? Evidence from tick data. Journal of Banking and Finance 33, 1741-1754.

Almeida, C., Vicente, J., 2009. Identifying volatility risk premia from fixed income Asian options. Journal of Banking and Finance 33, 652-661.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics. Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885-905.

Andersen, T.G., Bollerslev, T., Diebold, F., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of American Statistical Association 96, 42-55.

Andersen, T.G., Bollerslev, T., Diebold, F., Ebens, H., 2001. The distribution of realized stock return volatility. Journal of Financial Economics 61, 43-76.

Baillie, R., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3-30.

Baillie, R., Kapetanios, G., 2008. Semiparametric estimation of long memory: The Holy Grail or a poisoned chalice? Working paper.

Bandi, F., Perron, B., 2006. Long memory and the relation between implied and realized volatility. Journal of Financial Econometrics 4, 636-670.

Becker, R., Clements, A., White, S., 2007. Does implied volatility provide any information beyond that captured in model-based volatility forecasts? Journal of Banking and Finance 31, 2535-2549.

Breuer, J. B., Wohar, M.E., 1996. Overlapping and clumping problems with spot and forward exchange rate data. Economic Journal 106, 26-38.

Chernov, M., 2007. On the role of risk premia in volatility forecasting. Journal of Business and Economics Statistics 25, 411-426.

Christensen, B., Hansen, C., 2002. New evidence on the implied-volatility relation. European Journal of Finance 8, 187-205.

Christensen, B., Nielson, M., 2006. Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. Journal of Econometrics 133, 343-371.

Christensen B., Hansen, C., Prabhala, N., 2001. The telescoping overlap problem in options data. Working paper, University of Maryland.

Christensen, B., Prabhala, N., 1998. The relation between implied and realized volatility. Journal of Financial Economics 50, 125-150.

Davidson, J., 2002. A model of fractional cointegration, and tests for cointegration using the bootstrap. Journal of Econometrics 110, 187-212.

Davidson, J., 2005. Testing for fractional cointegration: The relationship between government popularity and economic performance in the UK. In: Diebolt, C., Kyrtsou, C. (Eds.), New Trends in Macroeconomics. Springer Verlag.

Dittmann, I., 2001. Fractional cointegration of voting and non-voting shares. Applied Financial Economics 11, 321-332.

Doornik, J., 1999. Object-oriented Matrix Programming using Ox. Timberlake Consultants Press: London.

Doran, J., Ronn, E., 2008. Computing the market price of volatility risk in the energy commodity markets. Journal of Banking and Finance 32, 2541-2552.

Dunis, C., Keller, A., 1995. Efficiency tests with overlapping data: An application to the currency options market. European Journal of Finance 4, 345-366.

Dunis, C., Huang, X., 2002. Forecasting and trading currency volatility: An application of recurrent neural regression and model combination. Journal of Forecasting 21, 317-354.

Gerolimetto, M., 2006. Frequency domain bootstrap for the fractional cointegration regression. Economics Letters 91, 389-394.

Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4, 221-238.

Granger, C., Joyeux, R., 1980. An introduction to long-memory models and fractional differencing. Journal of Time Series Analysis 1, 15-39.

Hannan, E.J., 1982. The estimation of the order of an ARMA process. Annals of Statistics 10, 1071-81.

Harvey, A., 1993. Time Series Models. Philip Allan: Deddington.

Hassler, U., Marmol, F., Velasco, C., 2006. Residual log-periodogram inference for long-run relationships. Journal of Econometrics 130, 165-207.

Hosking, J., 1981. Fractional differencing. Biometrika 68, 165-176.

Islam, A., Khaled, M., 2005. Tests of weak form efficiency of Dhaka stock market. Journal of Business Finance and Accounting 32, 1613-1624.

Kellard, N., 2006. On the robustness of cointegration tests for market efficiency. Finance Research Letters 3, 57-64.

Kellard, N., Sarantis, N., 2008. Can exchange rate volatility explain persistence in the forward premium? Journal of Empirical Finance 15, 714-728.

Kim, C., Phillips, P.C.B., 2001. Fully modified estimation of fractional cointegration models. Working paper, Yale University.

Konstantinidi, E., Skiadopoulos, G., Tzagkaraki, E., 2008. Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. Journal of Banking and Finance 32, 2401-2411.

Lobato, I., Nankervis, J., Savin, N., 2001. Testing for autocorrelation using a modified Box-Pierce Q test. International Economic Review 42, 187-205.

Lo, A. W., MacKinlay, A. C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41-66.

Lo, A. W., MacKinlay, A. C., 1989. The size and power of the variance ratio test in finite samples. Journal of Econometrics 40, 203-238.

Ljung, G., Box, G., 1978. On a measure of lack of fit in time series models. Biometrika 65, 297-303.

Marinucci, D., Robinson, P.M., 2001. Semiparametric fractional cointegration analysis. Journal of Econometrics 105, 225-247.

Maynard, A., Phillips, P.C.B., 2001. Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. Journal of Applied Econometrics 16, 671-708.

Neely, C., 2009. Forecasting foreign exchange volatility: Why is implied volatility biased and inefficient? And does it matter? Journal of International Financial Markets, Institutions and Money 19, 188-205.

Newbold, P., Rayner, A., Kellard, N., Ennew, C., 1998. Is the $/ECU exchange rate a random walk? Applied Financial Economics 8, 553-558.

Nielsen, M., 2007. Local whittle analysis of stationary fractional cointegration and the implied-realized volatility relation. Journal of Business and Economics Statistics 25, 427-446.

Ng, S., Perron, P., 1997. Estimation and inference in nearly unbalanced nearly cointegrated systems. Journal of Econometrics 79, 53-81.

Poteshman, A., 2000. Forecasting future volatility from option prices. Working paper, University of Illinois at Urbana-Champaign.

Pynnönen, S., Knif, J., 1998. Common long-term and short-term price memory in two Scandinavian stock markets. Applied Financial Economics 8, 257-265.

Robinson, P. M., 1995. Log-periodogram regression of time series with long range dependence. Annals of Statistics 23, 1048-1072.

Robinson, P.M., Marinucci, D., 2003. Semiparametric frequency domain analysis of fractional cointegration. In: Robinson, P.M. (Ed.), Time Series with Long Memory. Oxford University Press: Oxford.

Sarantis, N., 2006. Testing uncovered interest rate parity using traded volatility, a time varying risk premium and heterogeneous expectations. Journal of International Money and Finance 25, 1168-1186.

Tse, Y., Anh V., Tieng, A., 1999. No-cointegration test based on fractional differencing: Some Monte Carlo results. Journal of Statistical Planning and Inference 80, 257-267.

Velasco, C., 1999. Non-stationary log-periodogram regression. Journal of Econometrics, 91, 325-371.

Velasco, C., 2003. Gaussian semiparametric estimation of fractional cointegration. Journal of Time Series Analysis 24, 345-378.

Vilasuso, J., 2002. Forecasting exchange rate volatility. Economics Letters 76, 59-64.

Yang, J., Su, X., Kolari, J., 2008. Do Euro exchange rates follow a martingale? Some out-of-sample evidence. Journal of Banking and Finance 32, 729-740.

Table 1

1% Size of NBLS tests ($T = 250;\ z = T^{0.75};\ l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.022	0.027	0.029	0.021	0.023	0.021	0.021	0.021	0.021
0.7	-	0.024	0.025	0.021	0.024	0.019	0.021	0.018	0.016
0.6	-	-	0.019	0.017	0.020	0.017	0.019	0.019	0.015
0.5	-	-	-	0.015	0.015	0.017	0.018	0.014	0.014
0.4	-	-	-	-	0.012	0.015	0.017	0.016	0.014
0.3	-	-	-	-	-	0.015	0.016	0.016	0.016

5% Size of NBLS tests ($T = 250;\ z = T^{0.75};\ l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.073	0.074	0.074	0.073	0.076	0.074	0.069	0.066	0.064
0.7	-	0.066	0.070	0.067	0.066	0.065	0.066	0.060	0.060
0.6	-	-	0.064	0.067	0.063	0.062	0.059	0.059	0.056
0.5	-	-	-	0.055	0.057	0.061	0.056	0.059	0.052
0.4	-	-	-	-	0.054	0.056	0.051	0.057	0.055
0.3	-	-	-	-	-	0.052	0.050	0.051	0.053

10% Size of NBLS tests ($T = 250;\ z = T^{0.75};\ l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.127	0.124	0.123	0.126	0.126	0.128	0.127	0.122	0.115
0.7	-	0.118	0.122	0.126	0.120	0.117	0.121	0.116	0.112
0.6	-	-	0.115	0.120	0.118	0.114	0.108	0.113	0.102
0.5	-	-	-	0.113	0.119	0.113	0.108	0.106	0.102
0.4	-	-	-	-	0.107	0.109	0.102	0.103	0.096
0.3	-	-	-	-	-	0.102	0.099	0.098	0.090

Table 2

1% Size of NBLS tests ($T = 250;\; z = T^{0.3};\; l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.049	0.042	0.043	0.040	0.042	0.040	0.038	0.035	0.037
0.7	-	0.049	0.050	0.037	0.043	0.037	0.037	0.040	0.034
0.6	-	-	0.047	0.042	0.042	0.046	0.043	0.039	0.035
0.5	-	-	-	0.044	0.041	0.040	0.042	0.036	0.037
0.4	-	-	-	-	0.042	0.044	0.036	0.038	0.036
0.3	-	-	-	-	-	0.046	0.042	0.037	0.038

5% Size of NBLS tests ($T = 250;\; z = T^{0.3};\; l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.114	0.108	0.100	0.096	0.101	0.105	0.100	0.092	0.090
0.7	-	0.114	0.112	0.096	0.099	0.094	0.101	0.096	0.095
0.6	-	-	0.122	0.102	0.099	0.101	0.095	0.092	0.085
0.5	-	-	-	0.120	0.104	0.094	0.096	0.089	0.079
0.4	-	-	-	-	0.110	0.104	0.097	0.090	0.080
0.3	-	-	-	-	-	0.107	0.103	0.094	0.088

10% Size of NBLS tests ($T = 250;\; z = T^{0.3};\; l = 0$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.181	0.170	0.158	0.158	0.153	0.167	0.157	0.159	0.149
0.7	-	0.179	0.171	0.164	0.160	0.156	0.165	0.159	0.158
0.6	-	-	0.186	0.177	0.166	0.167	0.155	0.157	0.151
0.5	-	-	-	0.187	0.174	0.163	0.158	0.154	0.150
0.4	-	-	-	-	0.187	0.171	0.158	0.145	0.139
0.3	-	-	-	-	-	0.173	0.161	0.149	0.138

Table 3

1% Size of NBLS tests ($T = 250;\ \ z = T^{0.75};\ l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.014	0.018	0.018	0.016	0.017	0.017	0.015	0.014	0.013
0.7	-	0.019	0.016	0.016	0.018	0.016	0.015	0.013	0.013
0.6	-	-	0.017	0.018	0.018	0.018	0.012	0.013	0.011
0.5	-	-	-	0.019	0.017	0.015	0.011	0.015	0.011
0.4	-	-	-	-	0.015	0.016	0.013	0.012	0.012
0.3	-	-	-	-	-	0.017	0.015	0.013	0.013

5% Size of NBLS tests ($T = 250;\ \ z = T^{0.75};\ l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.067	0.063	0.059	0.061	0.055	0.056	0.055	0.050	0.047
0.7	-	0.064	0.061	0.061	0.059	0.054	0.058	0.054	0.050
0.6	-	-	0.061	0.064	0.059	0.057	0.057	0.053	0.053
0.5	-	-	-	0.062	0.062	0.059	0.056	0.055	0.054
0.4	-	-	-	-	0.056	0.058	0.061	0.057	0.050
0.3	-	-	-	-	-	0.055	0.059	0.055	0.051

10% Size of NBLS tests ($T = 250;\ \ z = T^{0.75};\ l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.114	0.120	0.109	0.110	0.105	0.106	0.108	0.102	0.102
0.7	-	0.118	0.116	0.108	0.106	0.110	0.112	0.106	0.103
0.6	-	-	0.113	0.111	0.102	0.103	0.111	0.113	0.104
0.5	-	-	-	0.111	0.102	0.110	0.115	0.112	0.101
0.4	-	-	-	-	0.102	0.107	0.113	0.109	0.105
0.3	-	-	-	-	-	0.101	0.110	0.108	0.107

Table 4

1% Size of NBLS tests ($T = 250$; $z = T^{0.3}$; $l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.015	0.015	0.017	0.017	0.018	0.017	0.016	0.017	0.015
0.7	-	0.021	0.018	0.015	0.014	0.014	0.015	0.016	0.014
0.6	-	-	0.020	0.016	0.018	0.018	0.016	0.016	0.013
0.5	-	-	-	0.016	0.018	0.016	0.016	0.015	0.015
0.4	-	-	-	-	0.016	0.015	0.012	0.014	0.013
0.3	-	-	-	-	-	0.017	0.016	0.012	0.012

5% Size of NBLS tests ($T = 250$; $z = T^{0.3}$; $l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.062	0.060	0.059	0.060	0.057	0.054	0.056	0.053	0.055
0.7	-	0.062	0.065	0.059	0.056	0.056	0.055	0.059	0.058
0.6	-	-	0.069	0.061	0.062	0.054	0.055	0.060	0.058
0.5	-	-	-	0.066	0.061	0.061	0.061	0.058	0.057
0.4	-	-	-	-	0.068	0.064	0.068	0.062	0.056
0.3	-	-	-	-	-	0.067	0.071	0.070	0.063

10% Size of NBLS tests ($T = 250$; $z = T^{0.3}$; $l = 1$)

$d \setminus \delta$	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0
0.8	0.108	0.106	0.107	0.107	0.103	0.101	0.107	0.104	0.105
0.7	-	0.113	0.115	0.110	0.111	0.099	0.101	0.105	0.111
0.6	-	-	0.115	0.109	0.107	0.107	0.105	0.106	0.108
0.5	-	-	-	0.118	0.119	0.113	0.111	0.107	0.103
0.4	-	-	-	-	0.123	0.126	0.120	0.114	0.109
0.3	-	-	-	-	-	0.173	0.161	0.149	0.138

Table 5

GPH tests for the d of individual volatility series

		$\hat{d}$	$(\hat{d}-1)/\sigma_d$	$\hat{d}/\sigma_d$
UK£ /US$	RV^h	0.546 (0.104)	-4.365	5.250
	RV^d	0.554 (0.104)	-4.288	5.327
	IV	0.622 (0.104)	-3.635	5.981
US$/Yen	RV^h	0.514 (0.104)	-4.673	4.942
	RV^d	0.303 (0.104)	-6.702	2.913
	IV	0.724 (0.104)	-2.654	6.962
US$/SF	RV^h	0.473 (0.104)	-5.067	4.548
	RV^d	0.424 (0.104)	-5.538	4.077
	IV	0.572 (0.104)	-4.115	5.500
Euro/Yen	RV^h	0.712 (0.139)	-2.072	5.122
	RV^d	0.592 (0.139)	-2.935	4.259
	IV	0.695 (0.139)	-2.194	5.000
Euro/UK£	RV^h	0.874 (0.139)	-0.906	6.288
	RV^d	0.716 (0.139)	-2.043	5.151
	IV	0.798 (0.139)	-1.453	5.741
Euro/US$	RV^h	0.770 (0.139)	-1.655	5.540
	RV^d	0.694 (0.139)	-2.201	4.993
	IV	0.851 (0.139)	-1.072	6.122

Note: numbers in parentheses alongside the estimates for d are the standard errors σ_d.

Table 6

Wald tests for the equality of the GPH estimates for RV and IV

	UK£ /US$	US$/SF	US$/Yen	Euro/Yen	Euro/UK£	Euro/US$
RV^h	0.168 [0.682]	0.320 [0.572]	1.162 [0.281]	0.006 [0.940]	0.129 [0.719]	0.131 [0.718]
RV^d	0.109 [0.741]	0.648 [0.421]	4.988 [0.026]	0.2198 [0.639]	0.159 [0.690]	0.485 [0.486]

Note: the Wald statistic has a $\chi^2(1)$ distribution. The figures in square brackets are *p* values.

Table 7

Robust estimation for the integration order of the (level) residuals in (1)

	UK£ /US$	US$/SF	US$/Yen	Euro/Yen	Euro/UK£	Euro/US$
RV^h	0.254 (0.119)	0.288 (0.119)	0.170 (0.119)	0.411 (0.166)	0.248 (0.166)	-0.117 (0.166)
RV^d	0.072 (0.119)	-0.061 (0.119)	0.003 (0.119)	0.151 (0.166)	-0.410 (0.166)	-0.218 (0.166)

Note: numbers in parentheses are standard errors.

Table 8

OLS estimates of (1)

		$\hat{\alpha}$	$\hat{\beta}$
UK£ /US$	RV^h	-0.791	0.646
	RV^d	-0.357	0.907
US$/SF	RV^h	-0.327	0.822
	RV^d	-0.432	0.841
US$/Yen	RV^h	-0.412	0.785
	RV^d	-0.518	0.814
Euro/Yen	RV^h	-0.213	1.131
	RV^d	0.207	0.878
Euro/UK£	RV^h	0.173	1.040
	RV^d	0.218	0.847
Euro/US$	RV^h	0.051	0.995
	RV^d	0.114	0.904

Table 9

NBLS estimates of (1)

		$\hat{\alpha}$	$\hat{\beta}$	95% CI for $\hat{\beta}$	90% CI for $\hat{\beta}$
UK£ /US$	RV^h	-0.654	0.702	[0.389, 1.035]	[0.434, 0.992]
	RV^d	-0.013	1.048	[0.827, 1.277]	[0.859, 1.247]
US$/SF	RV^h	-0.121	0.914	[0.684, 1.135]	[0.710, 1.108]
	RV^d	-0.106	0.987	[0.820, 1.151]	[0.848, 1.124]
US$/Yen	RV^h	-0.079	0.932	[0.799, 1.069]	[0.815, 1.051]
	RV^d	-0.194	0.957	[0.792, 1.130]	[0.817, 1.107]
Euro/Yen	RV^h	-0.441	1.230	[1.052, 1.403]	[1.065, 1.392]
	RV^d	-0.098	1.010	[0.905, 1.119]	[0.916, 1.102]
Euro/UK£	RV^h	-0.069	1.096	[0.868, 1.319]	[0.884, 1.308]
	RV^d	0.034	0.946	[0.865, 1.023]	[0.874, 1.014]
Euro/US$	RV^h	-0.101	1.062	[0.915, 1.210]	[0.929, 1.198]
	RV^d	-0.141	1.018	[0.843, 1.186]	[0.867, 1.169]

Table 10a

p-values of heteroscedasticity robust Ljung-Box tests for the high frequency forecast error between (RV^h, IV) pairs

M	UK£ /US$	US$/SF	US$/Yen	Euro/Yen	Euro/UK£	Euro/US$
1	0.000	0.000	0.487	0.011	0.000	0.269
2	0.000	0.000	0.513	0.001	0.000	0.488
3	0.000	0.000	0.073	0.000	0.000	0.325
4	0.000	0.000	0.129	0.000	0.000	0.469
5	0.000	0.000	0.066	0.000	0.000	0.492
6	0.000	0.000	0.070	0.000	0.000	0.368
7	0.000	0.000	0.109	0.000	0.000	0.437
8	0.000	0.000	0.149	0.000	0.000	0.515
9	0.000	0.000	0.150	0.000	0.000	0.429
10	0.000	0.000	0.189	0.000	0.000	0.416
11	0.000	0.000	0.086	0.000	0.000	0.459
12	0.000	0.000	0.121	0.000	0.000	0.467
13	0.000	0.000	0.163	0.000	0.000	0.535
14	0.000	0.000	0.209	0.000	0.000	0.585
15	0.000	0.000	0.158	0.000	0.000	0.574
16	0.000	0.000	0.186	0.000	0.000	0.638
17	0.000	0.000	0.226	0.000	0.000	0.695
18	0.000	0.000	0.240	0.000	0.000	0.747
19	0.000	0.000	0.218	0.000	0.000	0.779
20	0.000	0.000	0.263	0.000	0.000	0.809

Table 10b

p-values of heteroscedasticity robust Ljung-Box tests for the forecast error between (RV^d, IV) pairs

m	UK£ /US$	US$/SF	US$/Yen	Euro/Yen	Euro/UK£	Euro/US$
1	0.044	0.533	0.469	0.821	0.051	0.692
2	0.128	0.515	0.710	0.912	0.089	0.434
3	0.118	0.161	0.854	0.955	0.184	0.276
4	0.136	0.260	0.936	0.967	0.222	0.100
5	0.087	0.359	0.949	0.988	0.334	0.136
6	0.127	0.451	0.742	0.996	0.454	0.210
7	0.187	0.549	0.671	0.998	0.514	0.265
8	0.096	0.606	0.709	0.999	0.600	0.340
9	0.017	0.701	0.773	0.989	0.598	0.258
10	0.028	0.757	0.591	0.714	0.664	0.315
11	0.043	0.821	0.605	0.789	0.438	0.358
12	0.063	0.871	0.592	0.849	0.474	0.332
13	0.081	0.866	0.597	0.763	0.498	0.397
14	0.108	0.838	0.671	0.218	0.370	0.456
15	0.068	0.862	0.738	0.274	0.343	0.530
16	0.092	0.899	0.795	0.225	0.391	0.571
17	0.092	0.912	0.827	0.270	0.369	0.323
18	0.073	0.938	0.819	0.273	0.413	0.351
19	0.048	0.933	0.848	0.307	0.441	0.411
20	0.045	0.933	0.885	0.268	0.439	0.413