

The long memory of the forward premium

RICHARD T BAILLIE

Michigan State University, East Lansing, MI 48824, USA

AND

TIM BOLLERSLEV

Northwestern University and NBER,

J K Kellogg Graduate School of Management, Evanston, IL 60208, USA

The estimation of ARFIMA models by approximate maximum likelihood estimation methods, reveals the forward premia for the currencies of Canada, Germany and the UK *vis-à-vis* the US dollar, to be well described by a fractionally integrated process. These models imply that all the forward premia are mean reverting, although their autocorrelations are quite persistent. This degree of persistence has led other studies to erroneously conclude that the forward premia contains a unit root. (JEL C22, E31).

In a recent paper, Crowder (1994) has cast doubt on the validity of the widely accepted conventional wisdom that the forward premium is stationary. Similarly, Evans and Lewis (1993) and Mark *et al.* (1993) have argued that standard testing procedures suggest that forward premia are non-stationary. Many previous studies have concluded that spot and forward exchange rates are well described as martingale differences, so that their univariate autoregressive representations contain a unit root. At the same time several articles, *e.g.* Meese and Singleton (1982) and Baillie and Bollerslev (1989), have found spot and forward rates to be cointegrated with the estimated cointegrating vector parameters having an intercept close to zero and a slope coefficient close to one, so that the forward premium is stationary. These findings seem quite reasonable given that the forward premium is closely associated with risk, and that it is hard to see any theoretical reasons for a non-stationary risk premium.

The purpose of this paper is to demonstrate that the forward premium is indeed mean reverting, but in a rather special and possibly important way. Our explanation for Crowder's findings is simply that by the very nature of the tests he computed, he was forced to choose between an $I(0)$ and $I(1)$ process with which to describe the temporal dependence within the forward premium. On allowing for more general, fractionally integrated $I(d)$ processes, for $0 < d < 1$, we can reconcile his findings. Although the effect of an innovation is quite persistent, shocks to the forward premium eventually do die out at a slow hyperbolic rate of decay. Interestingly enough, there turns out to be remarkable similarities between the estimated ARFIMA models that best describe the forward

premium for the different currencies. This regularity, or stylized fact, deserves further attention, and may possibly indicate quite strong restrictions on models generating time dependent risk premium.

I. Autocorrelation properties of exchange rates

The following analysis uses the same data as Crowder (1994), which was originally provided by Data Resources Incorporated. Each exchange rate series is in terms of the number of US dollars per unit of foreign currency at the close of trading on the last business day of each month from January 1974 through December 1991, for a total of 215 observations.

Crowder (1994) considers the forward premium, $f_t - s_t$, where s_t is the logarithm of the spot exchange rate and f_t is the logarithm of the one month maturity forward rate. The application of standard Augmented Dickey–Fuller (ADF) unit root tests implies that it is not possible to reject a unit root in the Canadian and German forward premium, while the tests reveal mixed evidence on the existence of a unit root in the UK forward premium. Furthermore, the application of the Kwiatkowski, Phillips, Schmidt and Shin (1992), *i.e.* KPSS unit root test, which has a null hypothesis of stationarity, generally rejects the forward premium is $I(0)$. Hence the different tests strongly indicate a unit root in the forward premium.

However, a more careful interpretation of these tests is revealing. As noted by Lee and Schmidt (1993), the KPSS test is also quite powerful against fractionally integrated alternatives, so that although Crowder's KPSS test statistics are providing evidence against $I(0)$ behavior, this should not automatically be interpreted as being suggestive of an $I(1)$ process. Similarly, as demonstrated by Diebold and Rudebusch (1989b), the conventional ADF test for a unit root, or $I(1)$ behavior, has very low power against fractionally integrated alternatives.

Before considering the forward premium in detail, it is important to note that there is overwhelming evidence that the logarithm of the spot exchange rate, s_t , contains a unit root, *i.e.* is $I(1)$, while the approximate rate of return, $\Delta s_t \equiv s_t - s_{t-1}$, is not only $I(0)$, but also uncorrelated at higher lags. Perhaps the first article to report this finding was Cornell (1977), followed by numerous studies using more formal tests for unit roots, including Meese and Singleton (1982), Corbae and Ouliaris (1986), Baillie and Bollerslev (1989) and Baillie and Bollerslev (1993). One possible dissenting view is provided by Cheung (1993), who reports some evidence that exchange rates may be fractionally integrated with long memory. However, the grounds for believing that exchange rates are $I(1)$ are not only statistical. There are compelling economic reasons in terms of arbitrage conditions and weak form efficiency arguments. Furthermore, the intuition that returns should be $I(0)$ also justifies the $I(1)$ assumption of spot exchange rates. So far, any deviation from the stylized fact that nominal exchange rates are $I(1)$ with their corresponding returns being approximately uncorrelated appears marginal.

Figures 1 through 3 graph the first 24 autocorrelations for the logarithm of the spot exchange rate, s_t , the return on the rate, Δs_t , and the forward premium

$f_t - s_t$ for each of the three countries currencies. The correlograms of the spot rates and the returns are entirely consistent with the stylized facts discussed above. The autocorrelations of the forward premium exhibit considerable persistence out until two year lags. In Crowder's (1994) view the autocorrelations 'display the tell tale signs of non-stationarity, a smooth and slow rate of decay'. However,

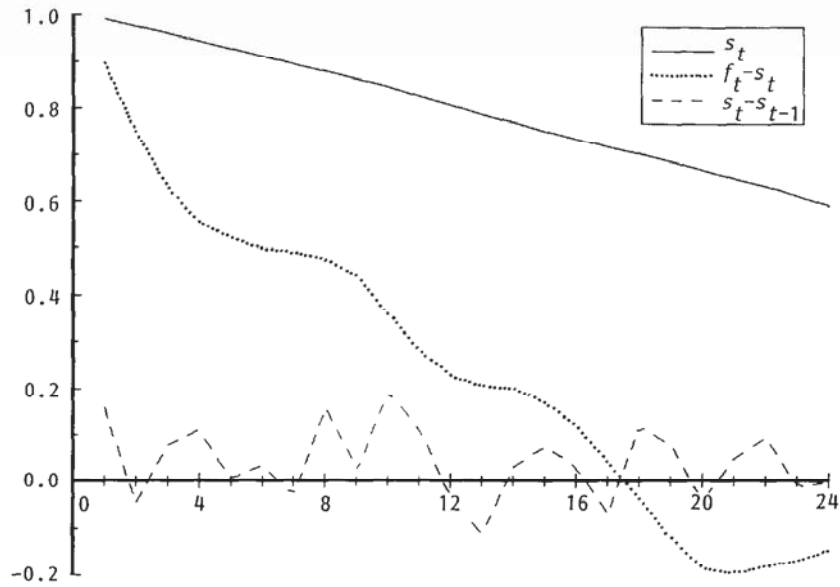


FIGURE 1. Canada.

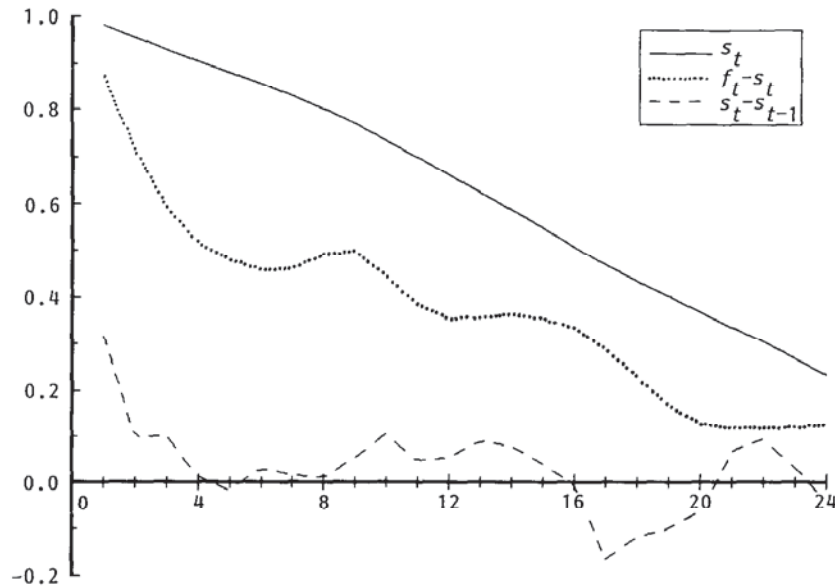


FIGURE 2. Germany.

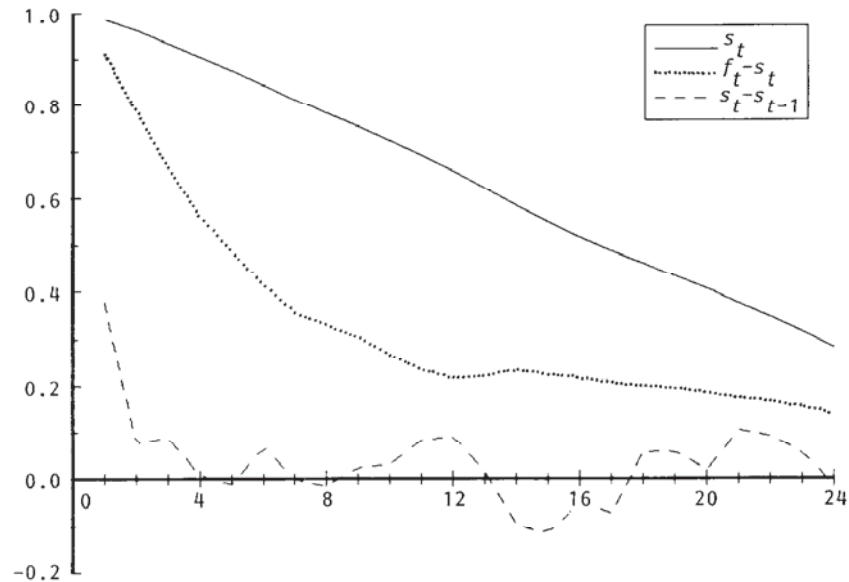


FIGURE 3. United Kingdom.

as is immediately evident from the three figures, the degree of persistence of the forward premium's autocorrelations are considerably less than those of the corresponding spot rates.

II. Fractionally integrated processes

Recently a number of researchers have expressed concern over the forced choice between $I(1)$ and $I(0)$ processes. For any covariance stationary time series process y_t , the impulse response weights or moving average representation are given by,

$$\langle 1 \rangle \quad y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

where the innovations ε_t are zero mean, finite variance and serially uncorrelated. This Wold decomposition only requires the square summability of $\sum_{j=0, \infty} \psi_j^2 < \infty$. However, the $I(0)$ property associated with stationary and invertible ARMA models imposes exponentially decaying impulse response weights on equation $\langle 1 \rangle$ and corresponding exponentially decaying autocorrelation coefficients. On the other hand, non-stationary $I(1)$ processes imply complete persistence on both the impulse response weights and the autocorrelation coefficients.

An important, more flexible class of processes, introduced by Granger and Joyeux (1980), Granger (1980, 1981) and Hosking (1981) is the ARFIMA(p, d, q) model,

$$\langle 2 \rangle \quad \phi(L)(1-L)^d y_t - \mu = \theta(L)\varepsilon_t,$$

where d denotes the fractional differencing parameter, all the roots of $\phi(L)$ and $\theta(L)$ lie outside the unit circle and ε_t is white noise. The fractional differencing operator, $(1-L)^d$ is defined in terms of its Maclaurins series expansion. The y_t process defined

by equation <2> and for $d \neq 0$, is then said to be $I(d)$. The Wold decomposition and autocorrelation coefficients will all exhibit a very slow rate of hyperbolic decay. For $-0.5 < d < 0.5$, the process is covariance stationary, while $d < 1$ implies mean reversion. In particular, for the fractional white noise processes,

$$(1 - L)^d y_t = \varepsilon_t,$$

the coefficients in the infinite moving average representation in equation <1> equals $\psi_j = \Gamma(j + d) / \{\Gamma(d)\Gamma(j + 1)\}$, where $\Gamma(\cdot)$ denotes the gamma function. The corresponding infinite autoregressive representation,

$$\langle 3 \rangle \quad y_t = \sum_{j=0}^{\infty} \pi_j y_{t-j} + \varepsilon_t,$$

has $\pi_j = \Gamma(j - d) / \{\Gamma(-d)\Gamma(j + 1)\}$. It follows by Sterling's formula that for high lags j , $\psi_j \approx \{\Gamma(d)\}^{-1} j^{d-1}$ and $\rho_j \approx \{\Gamma(1 - d) / \Gamma(d)\} j^{2d-1}$.

In this sense the $I(d)$ processes can be considered a half-way house between $I(0)$ and $I(1)$ processes. A more complete description of the general properties of these processes are given by Granger and Joyeux (1980) and Hosking (1981). Several recent studies, e.g. Diebold and Rudebusch (1989a), Diebold *et al.* (1991) and Sowell (1992a) have found the $I(d)$ class of processes to be useful in characterizing economic and financial time series, and have found it desirable to leave the $I(0)$ versus $I(1)$ paradigm.

In this paper an approximate time domain Maximum Likelihood Estimator, (MLE), that assumes conditional normality of ε_t is used to estimate ARFIMA models for the three forward premium series. The estimator is based on minimizing a Conditional Sum of Squares (CSS) function, and is asymptotically equivalent to full MLE under quite general conditional homoskedastic distributions. A description of the estimator together with detailed simulation evidence of its performance is presented by Chung and Baillie (1993). Specifically, all the parameter estimates are $T^{1/2}$ consistent, except the estimate of μ which converges at the slower rate of $T^{1/2-d}$. Also in moderate sized samples of 100 or more the CSS estimator is virtually identical in the sense of bias and mean squared error to the full MLE of Sowell (1992b).

TABLE 1. Estimation of ARFIMA(2,d,0) models for the forward premium.
 $[(1 - \phi_1 L - \phi_2 L^2)(1 - L)^d(1 - L)(s_t - f_t - \mu) = \varepsilon_t.]$

| Estimates | Canada | Germany | UK |
|------------|----------------|----------------|----------------|
| d | -0.555 (0.166) | -0.233 (0.094) | -0.449 (0.158) |
| ϕ_1 | 0.717 (0.170) | 0.489 (0.094) | 0.682 (0.170) |
| ϕ_2 | -0.202 (0.067) | -0.231 (0.068) | -0.074 (0.070) |
| μ | 0.001 (0.001) | -0.000 (0.000) | -0.002 (0.002) |
| σ^2 | 0.004 (0.000) | 0.008 (0.000) | 0.011 (0.001) |
| $Q(10)$ | 14.57 | 19.12 | 14.64 |

Key: All the models were estimated by minimizing the conditional sum of squares, which is approximately MLE under normality and is described in Section III of the text. The numbers in parentheses are the asymptotic standard errors of the corresponding parameter estimates.

III. Estimation and interpretation of results

The results of estimating ARFIMA(2, d ,0) models are reported in Table 1. The portmanteau tests for any remaining serial correlation in the residuals, $Q(10)$, indicate that this relatively simple model provides a good description of the own temporal dependencies in the conditional mean for all three forward premia.

Since all the models are estimated for the differenced series, $\Delta(f_t - s_t)$, the forward premium for Canada, Germany and the UK are really $I(0.45)$, $I(0.77)$ and $I(0.55)$ respectively. Hence the forward premia for Germany and the UK both have infinite variance, but finite cumulative impulse response weights, so that all three forward premia are estimated to be mean reverting. Indeed, the estimated models are all fairly similar in terms of the low order autoregressive structure and the values of d .

It is worth noting that estimation of the same models for the levels of the forward premia, $f_t - s_t$, gave rise to very similar results. Although it is possible to numerically estimate models with $d > 0.5$, or $d < -0.5$, formally the likelihood function is only defined for $-0.5 < d < 0.5$. However, for the sake of clarity we only report the results for the first differenced series for all three countries.

IV. Conclusion

It has become traditional to assume that the spot and forward series are cointegrated of the form $CI(1,1)$ in Engle and Granger's (1987) notation. Our results show that a better representation is obtained by allowing for fractional cointegration of the form $CI(1,d)$ with $0 < d < 1$, so that all forward premia may have infinite variances, but are nevertheless mean reverting with finite cumulative impulse response weights. There are many potential implications of this finding. From the theory of cointegrated processes, developed by Granger (1980), the forward market forecast error ($s_{t+1} - f_t$) will be of the same order of integration as the forward premium. This will therefore have implications for the appropriate distribution theory for the commonly employed regressions to test for unbiasedness of the forward rate, *e.g.* see the discussion in Hodrick (1987).

Similarly, the use of Δs_{t+1} and $f_t - s_t$ in a vector autoregression setting, as in Baillie (1989), also appears inappropriate since an error correction mechanism reflecting fractional cointegration will have to be included in the vector autoregression. Perhaps, most importantly the finding has interesting implications for measures of risk and the possible formulation of discrete time asset pricing models. These issues all seem worthy of further research.

References

- BAILLIE, R. T., 'Tests of rational expectations and market efficiency,' *Econometric Reviews*, 1989, **8**: 151-186.
- BAILLIE, R. T. AND T. BOLLERSLEV, 'Common stochastic trends in a system of exchange rates,' *Journal of Finance*, March 1989, **44**: 167-181.
- BAILLIE, R. T. AND T. BOLLERSLEV, 'Cointegration, fractional cointegration and exchange rate dynamics,' *Journal of Finance*, 1993, **49**: 737-745.
- CHEUNG, Y.-W., 'Long memory in foreign-exchange rates,' *Journal of Business and Economic Statistics*, January 1993, **11**: 93-101.

- CHUNG, C.-F. AND R. T. BAILLIE, 'Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models,' *Empirical Economics*, 1993, **18**, 791–806.
- CORBAE, D. AND S. OULIARIS, 'Robust tests for unit roots in the foreign exchange market,' *Economics Letters*, 1986, **22**: 375–380.
- CORNELL, B., 'Spot rates, forward rates and exchange market efficiency,' *Journal of Financial Economics*, 1977, **5**, 56–65.
- CROWDER, W. J., 'Foreign exchange market efficiency and common stochastic trends,' *Journal of International Money and Finance*, October 1994, (this issue).
- DIEBOLD, F. X., S. HUSTED AND M. RUSH, 'Real exchange rates under the gold standard,' *Journal of Political Economy*, December 1991, **99**: 1252–1271.
- DIEBOLD, F. X. AND G. D. RUDEBUSCH, 'Long memory and persistence in aggregate output,' *Journal of Monetary Economics*, September 1989, **24**: 189–209 (1989a).
- DIEBOLD, F. X. AND G. D. RUDEBUSCH, 'On the power of Dickey Fuller Tests against fractional alternatives,' *Economics Letters*, 1989, **35**: 155–160 (1989b).
- ENGLE, R. F. AND C. W. J. GRANGER, 'Cointegration and error correction: representation, estimation and testing,' *Econometrica*, March 1987, **55**: 251–276.
- EVANS, M. D. D. AND K. K. LEWIS, 'Do long term swings in the dollar affect estimates of the risk premia?,' working paper, University of Pennsylvania, 1993.
- GRANGER, C. W. J., 'Long memory relationships and the aggregation of dynamic models,' *Journal of Econometrics*, October 1980, **14**: 227–238.
- GRANGER, C. W. J., 'Some properties of time series data and their use in econometric model specification,' *Journal of Econometrics*, May 1981, **16**: 121–130.
- GRANGER, C. W. J. AND R. JOYEUX, 'An introduction to long memory time series models and fractional differencing,' *Journal of Time Series Analysis*, 1980, **1**: 15–39.
- HODRICK, R. J., *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, Harwood Academic Publishers, Chur, 1987.
- HOSKING, J. R. M., 'Fractional differencing,' *Biometrika*, 1981, **68**: 165–176.
- KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT AND Y. SHIN, 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?,' *Journal of Econometrics*, October–December 1992, **54**: 159–178.
- LEE, D. AND P. SCHMIDT, 'On the power of the KPSS test of stationarity against fractionally-integrated alternatives,' working paper, Michigan State University, 1993.
- MARK, N. C., Y. WU AND W. HAI, 'Understanding spot and forward exchange rate regressions,' working paper, Ohio State University, 1993.
- MEESE, R. A. AND K. J. SINGLETON, 'On unit roots and the empirical modeling of exchange rates,' *Journal of Finance*, September 1982, **37**: 1029–1035.
- SOWELL, F. B., 'Modeling long-memory behavior with the fractional ARMA model,' *Journal of Monetary Economics*, April 1992, **29**: 277–302 (1992a).
- SOWELL, F. B., 'Maximum likelihood estimation of stationary univariate fractionally integrated time series models,' *Journal of Econometrics*, July–September 1992, **53**, 165–188 (1992b).