European Economic Review 30 (1986) 365–381. North-Holland

# THE UNBIASEDNESS HYPOTHESIS IN THE FORWARD FOREIGN EXCHANGE MARKET

## A Specification Analysis with Application to France, Italy, Japan, the United Kingdom and West Germany

Allan W. GREGORY*

*University of Western Ontario, London, Ont., Canada, N6A 5C2*
*Queen's University, Kingston, Ont., Canada, K7L 3N6*

Thomas H. McCURDY*

*Queen's University, Kingston, Ont., Canada, K7L 3N6*

Received December 1984, final version received October 1985

The purpose of this paper is to carry out a specification analysis of a test relation for the unbiasedness hypothesis using thirty-day forward foreign exchange data from France, Italy, Japan, the United Kingdom, and West Germany. The results indicate that econometric problems do exist for each country's test equation over the entire sample. For each country, there is at least one period which admits a statistically adequate regression equation. Results for these periods show that the null hypothesis of the unbiasedness of the forward rate is rejected for some countries but is retained for others.

## 1. Introduction

In the last decade, there has been a plethora of empirical papers purporting to address some aspect of foreign exchange market efficiency. The consensus which seems to have emerged in this literature is that, in the absence of an explicit model, there is no definitive test for market efficiency. Instead researchers have preferred to test specific hypotheses (usually in the regression framework) and leave the more thorny issue of market efficiency in the background. Given the absence of a generally agreed upon empirical formulation of the market efficiency problem, this seems to be a sensible research strategy. At least this approach allows the stylized facts to be uncovered, and might also suggest appropriate models of the underlying market fundamentals.

*The authors would like to thank two referees for helpful comments. The second author acknowledges the financial support of the Social Sciences and Humanities Research Council of Canada.

0014–2921/86/$3.50 

Perhaps the most popular regression-based tests in this literature are those investigating the hypothesis of unbiasedness of the forward foreign exchange rate. Until very recently, the focus in application was upon parameter significance testing with very little concern for the adequacy of the test relation itself. That is, despite the fact that linear regression methods were used and statistical inferences were made, there was a surprising neglect of standard econometric practices. However, the direction of testing has now been extended from simple parameter significance testing to a wider consideration of the properties of the test relation as a whole. Excellent examples of this are Baillie, Lippens and McMahon (1983), Domowitz and Hakkio (1985), Hakkio (1981), and Hodrick and Srivastava (1984).

Evidence that the usual test relations of unbiasedness of the forward rate can be statistically inadequate is in Gregory and McCurdy (1984). In that investigation, we conducted a specification analysis of two regression-based test equations using thirty-day Canadian/American foreign exchange data. For that data set, we found that the test relations from full-sample estimation were extremely unreliable and consequently they could be very misleading. By construction, the null hypothesis of unbiasedness cannot be misspecified, but the statistical properties of the estimated test equation may be such that coefficients and/or standard errors are inconsistent – rendering parameter significance testing invalid.

Diagnostic tests in Gregory and McCurdy (1984) indicated that the 'ills' associated with the test relations were not the same over time. That is, the type of specification error committed depended upon the estimation period. Nevertheless, one striking feature for the Canadian/United States thirty-day foreign exchange market was that, in general, whenever the test relation 'passed' the battery of diagnostic tests applied, the hypothesis that the forward rate is an unbiased linear predictor of the future spot rate was supported by the data.

The purpose of this paper is to carry out a specification analysis of a test relation for the unbiasedness hypothesis using thirty-day forward foreign exchange data from France, Italy, Japan, United Kingdom and West Germany. In order to gain a broader understanding of the forward foreign exchange market, it is necessary to confirm or to contradict results from one data set with results from other data sets. Therefore this cross-country study compares the statistical properties of the test relations over the same time period. The data sets have been carefully constructed so that the timing of forward contracts and future spot rates are matched.

Our results show that a popular regression equation used to test the hypothesis that the forward premium is an unbiased linear predictor of the change in the corresponding spot rate, exhibits different kinds of econometric deficiences for different countries. In fact, the diagnostic tests indicate that estimation over the entire sample is always inappropriate using that test relation. Therefore, it is important to isolate statistically 'acceptable' specifications prior to parameter significance testing. There are occasions for which the simple regression equation is econometrically satisfactory. In those cases, it is possible to reject the parameter restrictions implied by the unbiasedness hypothesis for some countries – in contrast to the evidence in Gregory and McCurdy (1984) for the Canadian case.

The paper is organized as follows. In section 2 we review one test relation for unbiasedness, outline the specification tests and discuss the data construction. The results are presented in section 3 and concluding remarks are given in section 4.

## 2. Testing the unbiasedness hypothesis, specification tests and the data

### *2.1. Testing the unbiasedness hypothesis*

The null hypothesis that the forward premium is an unbiased linear predictor of the change in the corresponding spot rate may be tested using the regression equation:

$$DS_{t+1}^i = \alpha + \beta FP_t^i + u_{t+1}^i, \qquad \begin{array}{l} t=1,\dots,T, \\ i=1,\dots,N \text{ (countries)} \end{array} \tag{1}$$

in which

$$DS_{t+1}^i \equiv (S_{t+1}^i - S_t^i) \quad \text{and} \quad FP_t^i \equiv (F_t^i - S_t^i),$$

where $S_{t+1}$ is the logarithm of the spot exchange rate at time $t+1$, $F_t$ is the logarithm of the forward rate established at time $t$ for period $t+1$, and $u_{t+1}$ is an error term. Superscript $i$ indicates country $i$ and all exchange rates are quoted against U.S. dollars. In this form, the unbiasedness hypothesis implies that $\alpha=0$ and $\beta=1$. Such a restriction is consistent with a model of a competitive market with no transactions costs, risk-neutral speculators and market expectations which are rational. For that model, we would have

$$\mathrm{E}_t[DS_{t+1}^i] = FP_t^i, \tag{2}$$

where $\mathrm{E}_t$ is the mathematical expectation operation conditional upon some information set (and model).[1] The test relation (1) and the joint null hypothesis of rational expectations and no risk premium implicit in (2) can be related by decomposing the actual change in the spot rate into two

[1] Eq. (1) is not a structural model for determining the spot rate. The test is designed to investigate whether the conditional mean of the change in the future spot rate (measured in logarithms) is equal to the forward premium.

orthogonal components:

$$DS^i_{t+1} = \mathrm{E}_t[DS^i_{t+1}] + (DS^i_{t+1} - \mathrm{E}_t[DS^i_{t+1}]). \tag{3}$$

Substituting (2) into (3) yields (1) under the null. Testing the unbiasedness hypothesis involves estimating regression eq. (1) and determining whether the coefficient estimates of $\alpha$ and $\beta$ are significantly different from zero and one respectively.

The omission of transactions costs and the assumption of no risk premium in (2) implies that a rejection of the null ($\alpha=0$ and $\beta=1$) does not necessarily indicate a rejection of either the rational expectations hypothesis (REH) or the efficient market hypothesis (EMH). Since there is a substantial theoretical literature which demonstrates a time-varying risk premium, recent applied research has introduced additional variables in eq. (1) in an attempt to capture that premium [see, for example, Domowitz and Hakkio (1985), Hansen and Hodrick (1983), Hodrick and Srivastava (1984)]. Nevertheless, if the analysis is done in terms of regression eq. (1), the parameters $\alpha$ and $\beta$ must be time-invariant. We have argued in Gregory and McCurdy (1984) that such restrictions should be checked against the data. Certainly, omitted variables (such as a time-varying risk premium) and non-linearities could lead to instabilities in those parameter estimates.

Under the REH and EMH, the residuals in (1) should be serially independent.[2] Of course, misspecification of the systematic part of a regression equation could also manifest itself as serial correlation. In addition, estimation has often proceeded under the assumption that the errors are homoskedastic. Several authors [Cumby and Obstfeld (1983), Domowitz and Hakkio (1985), Hodrick and Srivastava (1984), and Hsieh (1984)] have observed that this is not implied by the unbiasedness hypothesis.

Applied research has frequently failed to pay sufficient attention to the properties of the residuals. Clearly, a careful research program would require that we conduct a thorough specification analysis of the test relation. That is, in order to evaluate an empirical specification properly, we need to examine and to test the regression equation in several directions. Prior to parameter significance testing, we should have some confidence that we indeed have a 'reasonable' specification.

### 2.2. *Specification tests*

With respect to test eq. (1), we have outlined two restrictions from the

[2]We note that there could be contemporaneous correlation between the $u^i_t$ (i.e., $\mathrm{E}\,u^i_t u^j_t \neq 0$). In the application we did not estimate the equations jointly but instead followed most of the literature and estimated each equation individually by ordinary least squares. Hence we might expect some loss of efficiency. Efficient estimates could be obtained using a SURE procedure as in Bailey, Baillie and McMahon (1984) and Bilson (1981).

unbiasedness hypothesis under rational expectations: (i) $\alpha=0$ and $\beta=1$, and (ii) $u_{t+1}^i$ is serially uncorrelated. These particular restrictions may be tested by estimating (1) by OLS, calculating a Wald test for $\alpha=0$ and $\beta=1$, and testing the least squares residuals for serial correlation using some Lagrange multiplier test. Notice that the direction of testing is mixed in the sense of being 'down' and 'up' from eq. (1). The downward direction involves testing whether the restrictions $\alpha=0$ and $\beta=1$ are supported by the data. In this case (1) is considered the 'unrestricted' equation. The validity of such a test is *conditional* upon eq. (1) being otherwise correctly specified.[3] On the other hand, the serial correlation test treats (1) as the 'restricted' equation and investigates the temporal independence of the errors assumed in the OLS estimation. This is testing in an upward direction. However, this is not the only upward direction that could be considered. Indeed there are many forms of specification error and it is useful to compare (1) against a variety of alternatives using several different diagnostic tests.

For the five countries (France, Italy, Japan, United Kingdom and West Germany) we conduct a detailed specification analysis of test eq. (1).[4] Inevitably, whenever a battery of diagnostic tests are applied to an estimated equation some confusion arises as to what the alternative hypothesis actually is. Certain tests, such as Ramsey's (1969) RESET test and the information matrix test of White (1982), do not yield well-defined alternative hypotheses. Diagnostic tests are applied sequentially with each test implying a different (perhaps unspecified) alternative hypothesis.

To complicate matters further, tests designed to detect a particular kind of specification error may nevertheless pick-up other kinds of statistical deficiencies [see Pagan and Hall (1983)]. For example, a test for $p$th order serial correlation for eq. (1) has the implied alternative

$$u_{t+1}^i = \sum_{j=1}^{p} \rho_j u_{t+1-j}^i + \varepsilon_{t+1}^i, \tag{4}$$

[3]Even if the alternative under which the test equation is estimated is misspecified, one may still have power against the null. However, the results are difficult to evaluate without a formal analysis of the relationship between the null and the alternative, plus an analysis of the power of estimators when estimating false models. See, for example, Davidson and MacKinnon (1985) and references therein.

[4]Orthogonality tests have also often been applied to determine whether $F_t^i$ is an optimal linear predictor of $S_{t+1}^i$. For these tests, one examines whether the forecast error, $S_{t+1}^i - F_t^i$ is orthogonal to some information set $I_t$. A test of unbiasedness is to test the null hypothesis that all regression coefficients are equal to zero. Following Gregory and McCurdy (1984) for each country, we regressed the forecast error upon its own lagged forecast error and its own forward premium:

$$S_{t+1}^i - F_t^i = \Phi_0 + \Phi_1(S_t^i - F_{t-1}^i) + \Phi_2(F_t^i - S_t^i) + w_{t+1}^i,$$

and conducted a specification analysis of this test relation over the same subperiods used in the text. Again there was compelling evidence for each country that the test equation was misspecified and that parameter significance testing over the entire sample could be highly misleading. These results are presented in table 2.

where $\varepsilon_{t+1}^{i}$ is serially independent. However, serial correlation tests such as in Godfrey (1978) will also have power against other alternatives such as omitted variables [Davidson and MacKinnon (1985) and Pagan and Hall (1983)] and autoregressive conditional heteroskedasticity [Engle, Hendry and Trumble (1985)]. Thus a significant test statistic for serial correlation does not lead to 'accepting' the alternative (4) that the errors of (1) are generated by a $p$th-order autoregressive process.

As Davidson and MacKinnon (1985), Godfrey (1984) and many others have argued, diagnostic testing may indicate the equation under consideration is misspecified but cannot determine why. After all, the alternative hypothesis is never under test. Hence discussion about possible alternative hypotheses (models) must always be speculative and tentative. With this in mind, we present the results of diagnostic testing without reference to alternative models. Possible sources of rejections will be considered in section 4.

The diagnostic tests which are calculated are based upon the Lagrange multiplier principle of testing. We estimate the test relation by ordinary least squares over several subperiods and then use these estimates to determine whether the data supports the statistical restrictions which have been implicitly made for consistent inference. For each subperiod, we calculate: (i) a test for fourth order serial correlation (AUTO) developed in Godfrey (1978); (ii) a test for fourth order serial correlation which is valid in the presence of heteroskedasticity ($\text{AUTO}_{\text{H}}$) derived in Domowitz and Hakkio (1983); (iii) a general test for heteroskedasticity (H) found in White (1980); (iv) a test for fourth order autoregressive conditional heteroskedasticity (ARCH) discussed in Engle (1982); (v) the Information matrix test (INFO) of White (1982) based upon a calculation suggested by Chester (1983); and (vi) Ramsey's (1969) RESET test using the square of the fitted values.[5] The information matrix test may be interpreted as a test for parameter constancy [Chesher (1984)], and the RESET test is a general test for model misspecification [Pagan and Hall (1983)]. Each of these large sample tests are compared against the chi-square distribution. We report a marginal significance level for each diagnostic test without making adjustments to take into account the fact that many tests are being jointly considered. The types of diagnostic tests that are applied here are not, in general, independent and consequently there is no easy way to determine the joint probabilities. We treat each diagnostic test as if it were calculated in isolation.

[5]Tests for first through third order AUTO, $\text{AUTO}_{\text{H}}$ and ARCH were also calculated (available from the authors upon request). For the most part, test statistics using these orders produced marginal significance levels similar to those based upon a fourth order process. In addition, we found that the cube of the fitted values for the RESET test could at times be important.

### 2.3. The data

The data set has been kindly made available by the Bank of Canada and has been used in a study by Longworth, Boothe and Clinton (1983). Since the timing and matching of forward contracts to the future spot rate are crucial in investigating the unbiasedness hypothesis, it is worthwhile to discuss the construction of the data set.

The monthly data was constructed from the Data Resources Inc. (DRI) data base of daily New York market opening bid and ask quotations for spot and one month forward rates over the period January 1974 to December 1981. All exchange rates are expressed as the number of U.S. dollars per unit of domestic currency. A single number for each rate was obtained by taking the average of the bid and ask price. The monthly observations were selected from the daily spot and forward rate series on a 'contract' basis. Accordingly, the rates were chosen to correspond to dates at which forward contracts opened and closed. The forward rates were from the last Tuesday of each month in order to minimize 'weekend' effects [Levi (1978)]. The spot rates were chosen to match the forward rate contract value dates. The value dates were identified according to rules outlined in Longworth, Boothe and Clinton (1983, pp. 57–58). Their procedure 'guarantees that the chosen spot rates correspond exactly to the rates at which the forward contracts made on the last Tuesday of each month would close and represent actual prices faced by market participants'.[6]

For each country we estimate the test relation and calculate the diagnostic tests over the full sample (1974–1981) and over three overlapping subperiods (1974–1977, 1976–1979, and 1978–1981). These arbitrary divisions for subperiods are intended to illustrate the different and changing econometric structure of the test relation for the different countries. Our goal is to draw attention to the fact that econometric problems do exist for all the countries considered using the simple test eq. (1).

## 3. The results

We commence with a discussion of the results for each of the five countries in turn. These results appear in table 1.

[6]Although forward and future spot rates are exactly matched according to the appropriate contract dates, there may occasionally be a short overlap of contracts. This is because the forward rate is from the last Tuesday of each month which can precede the end-of-contract date corresponding to the previous month's forward rate. The potential serial correlation of the residuals due to this occasional overlap of forecast periods is not likely to be a problem since our test results indicate that whenever serial correlation is present, the process appears to be greater than order one. We would like to thank a referee for raising this issue.

Table 1

Testing the unbiasedness hypothesis: Cross-country evidence.[a]

| Sample period | $\alpha$ | $\beta$ | W | AUTO | $\text{AUTO}_H$ | H | ARCH | INFO | RESET | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| *France* | | | | | | | | | | |
| 1974–1981 | −0.0050 | −1.904 | 0.002 | 0.308 | 0.460 | 0.056 | 0.104 | 0.098 | 0.215 | 0.054 |
| | (0.0035) | (0.83) | | | | | | | | |
| 1974–1977 | −0.0068 | −2.216 | 0.046 | 0.854 | 0.595 | 0.815 | 0.768 | 0.779 | 0.151 | 0.033 |
| | (0.0079) | (1.76) | | | | | | | | |
| 1976–1979 | 0.0037 | 0.243 | 0.301 | 0.027 | 0.006 | 0.095 | 0.036 | 0.024 | 0.624 | 0.0008 |
| | (0.0046) | (1.30) | | | | | | | | |
| 1978–1981 | −0.0044 | −1.953 | 0.046 | 0.059 | 0.006 | 0.210 | 0.0008 | 0.175 | 0.129 | 0.05 |
| | (0.0049) | (1.26) | | | | | | | | |
| *Italy* | | | | | | | | | | |
| 1974–1981 | −0.0109 | −0.595 | 0.002 | 0.203 | 0.148 | 0.177 | 0.025 | 0.241 | 0.177 | 0.018 |
| | (0.0042) | (0.45) | | | | | | | | |
| 1974–1977 | −0.0113 | −0.569 | 0.009 | 0.298 | 0.142 | 0.563 | 0.007 | 0.392 | 0.133 | 0.023 |
| | (0.0063) | (0.54) | | | | | | | | |
| 1976–1979 | 0.0003 | 0.203 | 0.033 | 0.923 | 0.851 | 0.976 | 0.089 | 0.992 | 0.92 | 0.004 |
| | (0.0052) | (0.47) | | | | | | | | |
| 1978–1981 | −0.0113 | −0.836 | 0.145 | 0.364 | 0.075 | 0.331 | 0.530 | 0.572 | 0.396 | 0.015 |
| | (0.0063) | (0.99) | | | | | | | | |
| *Japan* | | | | | | | | | | |
| 1974–1981 | 0.0032 | 0.010 | 0.124 | 0.195 | 0.338 | 0.006 | 0.071 | 0.0001 | 0.752 | $3 \times 10^{-6}$ |
| | (0.0033) | (0.49) | | | | | | | | |
| 1974–1977 | 0.0061 | 0.823 | 0.087 | 0.035 | 0.006 | 0.147 | 0.034 | 0.231 | 0.045 | 0.08 |
| | (0.0029) | (0.40) | | | | | | | | |
| 1976–1979 | 0.0138 | −2.173 | 0.073 | 0.402 | 0.588 | 0.063 | 0.128 | 0.0009 | 0.027 | 0.04 |
| | (0.0066) | (1.51) | | | | | | | | |
| 1978–1981 | 0.0162 | −2.777 | 0.061 | 0.785 | 0.716 | 0.887 | 0.629 | 0.896 | 0.254 | 0.056 |
| | (0.011) | (1.70) | | | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *United Kingdom* | | | | | | | | | | |
| 1974–1981 | −0.0049<br>(0.0037) | −1.759<br>(0.85) | 0.004 | 0.123 | 0.097 | 0.110 | 0.104 | 0.037 | 0.327 | 0.044 |
| 1974–1977 | −0.0042<br>(0.0058) | −0.566<br>(1.08) | 0.259 | 0.001 | $6\times10^{-7}$ | 0.174 | 0.058 | 0.037 | 0.0001 | 0.006 |
| 1976–1979 | 0.0068<br>(0.0064) | 0.629<br>(1.41) | 0.199 | 0.242 | 0.042 | 0.681 | 0.672 | 0.741 | 0.093 | 0.004 |
| 1978–1981 | −0.0035<br>(0.0049) | −4.893<br>(1.57) | 0.0007 | 0.620 | 0.304 | 0.607 | 0.62 | 0.615 | 0.087 | 0.175 |
| *West Germany* | | | | | | | | | | |
| 1974–1981 | 0.00765<br>(0.0053) | −1.884<br>(1.47) | 0.142 | 0.717 | 0.637 | 0.018 | 0.016 | 0.028 | 0.979 | 0.017 |
| 1974–1977 | 0.0116<br>(0.0055) | −3.934<br>(3.02) | 0.109 | 0.785 | 0.496 | 0.931 | 0.87 | 0.919 | 0.56 | 0.026 |
| 1976–1979 | 0.0090<br>(0.008) | 0.328<br>(2.27) | 0.257 | 0.179 | 0.031 | 0.058 | 0.012 | 0.001 | 0.446 | 0.0005 |
| 1978–1981 | −0.0005<br>(0.013) | −0.279<br>(2.73) | 0.487 | 0.837 | 0.677 | 0.161 | 0.044 | 0.026 | 0.430 | 0.0002 |

[a]Standard errors are given in parentheses. W is the Wald test of the joint hypothesis that $\alpha=0$ and $\beta=1$, AUTO is a test for fourth order serial correlation [Godfrey (1978)], $\text{AUTO}_H$ is a heteroskedasticity-robust test for fourth order serial correlation [Domowitz and Hakkio (1983)], H is a general test of heteroskedasticity [White (1980)], ARCH is a test for fourth order conditional heteroskedasticity [Engle (1982)], INFO is the information matrix test [White (1982)] based upon the calculation of Chesher (1983), RESET is Ramsey's (1969) misspecification test using the square of the fitted values and $R^2$ is the coefficient of determination. Marginal significance levels are reported for each statistic.

Table 2

Orthogonality tests: Cross-country evidence.[a]

| Sample period | $\Phi_0$ | $\Phi_1$ | $\Phi_2$ | W | AUTO | $AUTO_H$ | H | ARCH | INFO | RESET | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *France* | | | | | | | | | | | |
| 1974–1981 | −0.0053 | −0.0446 | −2.731 | 0.0002 | 0.218 | 0.339 | 0.193 | 0.050 | 0.083 | 0.040 | 0.18 |
| | (0.0034) | (0.017) | (0.80) | | | | | | | | |
| 1974–1977 | −0.0041 | −0.042 | −2.241 | 0.001 | 0.681 | 0.184 | 0.973 | 0.732 | 0.593 | 0.535 | 0.22 |
| | (0.0074) | (0.014) | (1.66) | | | | | | | | |
| 1976–1979 | 0.0037 | −0.0042 | −0.723 | 0.492 | 0.031 | 0.006 | 0.036 | 0.035 | 0.053 | 0.082 | 0.009 |
| | (0.0046) | (0.018) | (1.32) | | | | | | | | |
| 1978–1981 | −0.0044 | 0.0004 | −2.955 | 0.110 | 0.058 | 0.005 | 0.552 | 0.0008 | 0.431 | 0.125 | 0.11 |
| | (0.0051) | (0.022) | (1.27) | | | | | | | | |
| *Italy* | | | | | | | | | | | |
| 1974–1981 | −0.0112 | −0.0053 | −1.591 | 0.003 | 0.169 | 0.102 | 0.158 | 0.028 | 0.041 | 0.230 | 0.13 |
| | (0.0042) | (0.0042) | (0.45) | | | | | | | | |
| 1974–1977 | −0.0123 | −0.0054 | −1.595 | 0.009 | 0.287 | 0.093 | 0.080 | 0.011 | 0.001 | 0.171 | 0.19 |
| | (0.0063) | (0.0039) | (0.54) | | | | | | | | |
| 1976–1979 | 0.0026 | 0.0103 | −0.683 | 0.006 | 0.157 | 0.345 | 0.067 | 0.123 | 0.072 | 0.862 | 0.22 |
| | (0.0048) | (0.0034) | (0.44) | | | | | | | | |
| 1978–1981 | −0.0117 | −0.0034 | −1.818 | 0.220 | 0.335 | 0.059 | 0.545 | 0.515 | 0.292 | 0.446 | 0.082 |
| | (0.0063) | (0.0044) | (0.99) | | | | | | | | |
| *Japan* | | | | | | | | | | | |
| 1974–1981 | 0.0028 | −0.0039 | −0.906 | 0.202 | 0.195 | 0.340 | 0.010 | 0.065 | 0.0001 | 0.823 | 0.046 |
| | (0.0034) | (0.0056) | (0.51) | | | | | | | | |
| 1974–1977 | 0.0056 | −0.0057 | −0.013 | 0.055 | 0.057 | 0.017 | 0.371 | 0.027 | 0.125 | 0.015 | 0.058 |
| | (0.0029) | (0.0036) | (0.41) | | | | | | | | |
| 1976–1979 | 0.0142 | 0.0017 | −3.226 | 0.157 | 0.409 | 0.599 | 0.055 | 0.131 | 0.001 | 0.028 | 0.09 |
| | (0.0068) | (0.0059) | (1.53) | | | | | | | | |
| 1978–1981 | 0.016 | −0.0010 | −3.760 | 0.139 | 0.794 | 0.716 | 0.740 | 0.629 | 0.570 | 0.254 | 0.098 |
| | (0.011) | (0.007) | (1.72) | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *United Kingdom* | | | | | | | | | | | |
| 1974–1981 | −0.0043 | 0.0563 | −2.316 | 0.003 | 0.269 | 0.181 | 0.087 | 0.121 | 0.017 | 0.147 | 0.12 |
| | (0.0037) | (0.036) | (0.89) | | | | | | | | |
| 1974–1977 | −0.0020 | 0.0763 | −0.745 | 0.019 | 0.014 | $2\times10^{-6}$ | 0.293 | 0.157 | 0.010 | 0.002 | 0.17 |
| | (0.0056) | (0.029) | (1.07) | | | | | | | | |
| 1976–1979 | 0.0067 | 0.0082 | −0.365 | 0.362 | 0.238 | 0.050 | 0.931 | 0.663 | 0.644 | 0.340 | 0.23 |
| | (0.0065) | (0.044) | (1.43) | | | | | | | | |
| 1978–1981 | −0.0033 | −0.0224 | −6.004 | 0.002 | 0.595 | 0.262 | 0.697 | 0.645 | 0.531 | 0.075 | 0.24 |
| | (0.0049) | (0.046) | (1.60) | | | | | | | | |
| *West Germany* | | | | | | | | | | | |
| 1974–1981 | 0.0067 | −0.045 | −2.722 | 0.100 | 0.681 | 0.675 | 0.022 | 0.024 | 0.039 | 0.590 | 0.063 |
| | (0.0053) | (0.03) | (1.46) | | | | | | | | |
| 1974–1977 | 0.0097 | −0.037 | −3.989 | 0.017 | 0.812 | 0.681 | 0.954 | 0.923 | 0.210 | 0.413 | 0.10 |
| | (0.0056) | (0.024) | (3.05) | | | | | | | | |
| 1976–1979 | 0.0081 | −0.025 | −0.479 | 0.338 | 0.250 | 0.033 | 0.002 | 0.01 | 0.006 | 0.006 | 0.016 |
| | (0.0082) | (0.03) | (2.29) | | | | | | | | |
| 1978–1981 | −0.0016 | −0.060 | −1.306 | 0.396 | 0.769 | 0.614 | 0.175 | 0.059 | 0.082 | 0.446 | 0.037 |
| | (0.013) | (0.05) | (2.71) | | | | | | | | |

[a]Standard errors are given in parentheses. W is the Wald test of the joint hypothesis that $\Phi_0=\Phi_1=\Phi_2=0$, AUTO is a test for fourth order serial correlation [Godfrey (1978)], $\mathrm{AUTO_H}$ is a heteroskedasticity-robust test for fourth order serial correlation [Domowitz and Hakkio (1983)], H is a general test of heteroskedasticity [White (1980)], ARCH is a test for fourth order conditional heteroskedasticity [Engle (1982)], INFO is the information matrix test [White (1982)] based upon the calculation of Chesher (1983), RESET is Ramsey's (1969) misspecification test using the square of the fitted values and $R^2$ is the coefficient of determination. Marginal significance levels are reported for each statistic.

### 3.1. France

On the basis of full-sample estimation (1974–1981), the joint null hypothesis that $\alpha=0$ and $\beta=1$ is rejected at the five percent level of confidence. Subsample estimation indicates that for two out of the three subperiods this rejection is confirmed. The null hypothesis is retained for the subperiod 1976–1979; however, with the exception of the RESET test, all of the diagnostic tests agree that the test relation is misspecified. According to both tests (AUTO and $\mathrm{AUTO_H}$), serial correlation is also present for the 1978–1981 period. At times, some heteroskedasticity is detected by both the H and the ARCH tests. Although the two tests are designed to detect heteroskedasticity, they can in practice give conflicting signals. For example, in the 1978–1981 period, we retain the hypothesis of homoskedastic errors for the H test but reject it using the ARCH test. Evidently, the postulated form of the heteroskedasticity is important in specification tests. The INFO test clearly rejects the specification for the subperiod 1976–1979. However, using the full sample the INFO test only rejects at a level of significance of ten percent. The null hypothesis under the RESET test is not rejected for any of the estimated periods. It is interesting to note that for the 1974–1977 period there is no strong evidence of misspecification and that the unbiasedness hypothesis is rejected (at the five percent level). This result contrasts to that obtained from Canadian/United States data for which there was a general tendency to retain the unbiasedness hypothesis whenever the estimating equation was deemed satisfactory [Gregory and McCurdy (1984)].

### 3.2. Italy

In contrast to the results obtained for France, the Italian test equation appears to be somewhat better behaved. For example, unlike the French equation, the INFO test indicates that the parameter estimates are fairly stable over time. The estimating equation is statistically 'acceptable' for the period 1976–1981. Interestingly, the unbiasedness hypothesis is rejected for the 1976–1979 period, but is retained for the 1978–1981 period. For the remaining periods, the ARCH test rejects the hypothesis that the errors are homoskedastic, but the other diagnostic tests register no additional statistical difficulties. While the estimated coefficients of $\alpha$ can be both positive and negative, it is only significant when it is negative.

### 3.3. Japan

The Japanese test equation displays a variety of econometric deficiencies. The Wald test for $\alpha=0$ and $\beta=1$ is retained for the full sample and all the subperiods (at the five percent level). However, from a statistical point of view, only the period 1978–1981 produces a reasonable empirical specification. The full sample results show heteroskedasticity (both H and ARCH) and a rejection according to the INFO test. Recall that this latter test can be interpreted as a test of parameter constancy [Chesher (1984)]. Such an interpretation seems well-supported as the coefficient estimates of $\beta$ are extremely unstable ranging from negative to positive significant. Serial correlation and ARCH errors are detected for the period 1974–1977. The RESET test also indicates misspecification. The INFO and RESET tests indicate substantial problems with the specification for the 1976–1979 subperiod.

### *3.4. United Kingdom*

As has been the case for France, Italy and Japan, there is at least one period for which the estimated regression equation 'passes' all of the diagnostic tests (at the five percent level). This is 1978–1981 in the U.K. case and the unbiasedness hypothesis is strongly rejected for that period. The remaining subsamples (and full sample) show evidence of some serial correlation (AUTO or $\mathrm{AUTO_H}$) or reject the null for the INFO and RESET tests. The parameters appear particularly unstable with the estimated coefficient of $\beta$ ranging from a large negative significant value to a positive insignificant one. The estimated coefficient of $\alpha$ can either be negative or positive but is never significant. As an indication of how incomplete parameter significance testing can be, consider the time period 1974–1977. The null hypothesis of unbiasedness is retained and yet there is ample evidence of misspecification (AUTO, $\mathrm{AUTO_H}$, INFO and RESET). Obviously in such a situation the results from parameter significance testing can be highly misleading. Finally, contrary to the results thus far, heteroskedasticity does not seem to pose a problem for the United Kindom data.

### *3.5. West Germany*

For all subsamples as well as the full sample we are unable to reject the hypothesis of unbiasedness (at the five percent level). Unlike the United Kingdom results, the estimated equation is statistically appropriate for the period 1974–1977, but again it is the only one. The specification for the period 1976–1979 is rejected by $\mathrm{AUTO_H}$, H, ARCH and INFO. Heteroskedasticity is present in all periods except for 1974–1977. The estimates of $\beta$ are highly variable over the different subperiods, and the coefficient estimate of $\alpha$ is significant only when it is positive (compare the Italian case).

## 4. Concluding remarks

Unbiasedness is a joint hypothesis which embodies several components including a zero risk premium, market expectations which are rational, and competitive markets with negligible transactions costs. This cross-country study examines and compares the statistical properties of a popular regression equation for testing the unbiasedness hypothesis in the thirty-day forward foreign exchange market. Diagnostic tests on each country's test equation indicate that estimation over the entire sample period is inappropriate. There is strong evidence to suggest that there are structural instabilities in both the systematic and stochastic components. Therefore, to avoid misleading conclusions, it is important to isolate statistically 'acceptable' specifications prior to parameter significance testing.

For each country, there is at least one subperiod which admits a statistically adequate regression equation. For these periods, the French and United Kingdom data always reject the unbiasedness hypothesis, whereas the null is retained for the Japanese and West German cases. The Italian results are conflicting in that the null hypothesis is rejected for the 1976–1979 period but is retained for the 1978–1981 period.

Although we have found that for some countries there are occasions in which specification (1) is satisfactory, the overall impression is that, as it stands, it is statistically inadequate. Determining the precise source of the rejection is likely to prove very difficult. As is well known, tests for misspecification can only detect an inadequate specification and thus cannot be expected to provide a reliable guide to re-specification. However, there have been some recent studies which have investigated alternatives which might shed some light on the problem. Common to each of these is the recognition that specification (1) is incomplete.

Possible alternatives considered in the literature include: (i) the error terms are heteroskedastic, (ii) the existence of a time-varying risk premium, (iii) the 'peso problem', and (iv) unexploited profit opportunities. We will briefly discuss these alternative 'explanations'. However, it should be kept in mind that this discussion is intended to stimulate further inquiry, and should not be interpreted as conclusions following from the empirical results presented above. Judgements about alternative models must be suspended until empirical analysis of the specific alternatives is undertaken.

One possible source of the statistical inadequacy of the test relation (1) is the presence of heteroskedasticity. Under the unbiasedness hypothesis, the errors need not be homoskedastic so that the least squares estimate of the covariance matrix may be inconsistent. It is unlikely that all the misspecification reported above is due to ignored heteroskedasticity. Tests robust for heteroskedasticity reveal that there is also some serial correlation present in the residuals which constitutes evidence against the null. Furthermore, studies by Gregory (1985) and Hansen and Hodrick (1980 and 1983), using heteroskedastic-consistent covariance matrices, indicate a rejection of the parameter restrictions implied by unbiasedness for test equations like (1).

Perhaps the most frequent explanation for the rejection of unbiasedness [and the presence of specification errors in eq. (1)] is the existence of a time-varying risk premium. In that case, estimating eq. (1) could result in significant statistics for many of the diagnostic tests which we have applied. A popular current research strategy has been to include, in (1), variables which purport to capture some element of a time-varying risk premium [see, for example, Domowitz and Hakkio (1985), Hansen and Hodrick (1983), Hodrick and Srivastava (1984), and Korajczyk (1985)]. Overall, the empirical evidence presented thus far suggests that the risk premia which have been modelled may not account for all the misspecification in eq. (1). Nevertheless, incorporating additional features of the risk premia may be worthwhile.

The 'peso problem' or rational expectations of a drastic policy regime shift, such as a major devaluation which does not materialize during the sample period, is another possible source of rejection of the unbiasedness hypothesis. While there is some discussion of this phenomenon in the literature [see, for example, Krasker (1980)], there has been very little empirical testing to verify the quantitative importance of the 'peso problem'. It is possible that the INFO test results presented above are detecting structural shifts due to this problem.

Finally, the existence of some inefficiency in the forward market that leads to unexploited profit opportunities could introduce serial correlation (last period's forward premium prediction error is correlated with that of this period) and thus imply a rejection of the unbiasedness hypothesis. Research into this alternative has included the out-of-sample risk-return tradeoff of filter rules [Dooley and Shafer (1976), (1983)], and the profitability of trading strategies [Bilson (1981), Hodrick and Srivastava (1984), and Longworth, Boothe and Clinton (1983)].

Certainly, the presence of misspecification implies the need for an alternative hypothesis derived from a (possibly more general) stable model. Nevertheless, the unbiasedness hypothesis per se is an interesting hypothesis – especially for macro model builders who require proxies for spot exchange rate expectations. In those cases for which (1) is an 'acceptable' test relation, we were able to reject unbiasedness for some countries but not for others. This suggests that country-specific models might be required. For example, in terms of underlying economic structure, the behaviour of the monetary authority could be an important determinant of a time-varying risk premium. However, those rejections of unbiasedness could alternatively be the result of cross-country differences in speculative behaviour and/or institutional features. Separating the components of the joint hypotheses inherent in the unbiasedness hypothesis remains an important direction for research.

## References

Bailey, Ralph W., Richard T. Baillie and Patrick C. McMahon, 1984, Interpreting econometric evidence of efficiency in the foreign exchange market, Oxford Economic Papers 36, 67–85.

Baillie, Richard T., Robert E. Lippens and Patrick C. McMahon, 1983, Testing rational expectations and efficiency in the foreign exchange market, Econometrica 51, 553–563.
Bilson, J.F.O., 1981, The speculative efficiency hypothesis, Journal of Business 54, 435–452.
Chesher, Andrew, 1983, The information matrix test: Simplified calculation via a score test interpretation, Economics Letters 13, 45–48.
Chesher, Andrew, 1984, Testing for neglected heterogeneity, Econometrica 52, 865–872.
Cumby, Robert E. and Maurice Obstfeld, 1983, International interest-rate and price level linkages under flexible exchange rates: A review of recent evidence, in: J.F.O. Bilson and R.C. Marston, eds., Exchange Rates: Theory and Practice (University of Chicago Press for the NBER, Chicago, IL).
Davidson, Russell and James MacKinnon, 1985, On the interpretation of test statistics, Canadian Journal of Economics 18, 38–57.
Domowitz, Ian and Craig S. Hakkio, 1983, Testing for serial correlation in the presence of heteroskedasticity with applications to exchange rate models, unpublished manuscript, Oct., 31 pp. (Northwestern University, Evanston, IL).
Domowitz, Ian and Craig S. Hakkio 1985, Conditional variance and the risk premium in the foreign exchange market, Journal of International Economics 19, 47–66.
Dooley, M.P. and J.R. Shafer, 1976, Analysis of short-run exchange rate behaviour: March, 1973 to September, 1975, International Finance Discussion Paper 123 (Federal Reserve Board, Washington, DC).
Dooley, M.P. and J.R. Shafer, 1983, Analysis of short-run exchange rate behaviour: March, 1973 to November, 1981, in: D. Bigman and T. Taya, eds., Exchange rate and trade instability: Causes, consequences, and remedies (Ballinger, Cambridge, MA).
Engle, Robert F. 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflations, Econometrica 50, 987–1007.
Engle, Robert F., David F. Hendry and David Trumble 1985, Small-sample properties of ARCH estimators and tests, Canadian Journal of Economics 18, 66–93.
Godfrey, Les G. 1978, Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables, Econometrica 46, 1303–1310.
Godfrey, Les G. 1984, On the use of misspecification checks and tests of the non-nested hypotheses in empirical econometrics, Economic Journal, Supplement to vol. 94, 69–81.
Gregory, Allan W. 1985, Estimating equations with combined moving average error processes under rational expectations, unpublished manuscript (Queen's University, Kingston).
Gregory, Allan W. and Thomas H. McCurdy 1984, Testing the unbiasedness hypothesis in the forward foreign exchange market: A specification analysis, Journal of International Money and Finance 3, 357–368.
Hakkio, Craig S. 1981, Expectations and the forward exchange rate, International Economic Review 22, 663–678.
Hansen, Lars Peter and Robert J. Hodrick 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, Journal of Political Economy 88, 829–853.
Hansen, Lars Peter and Robert J. Hodrick 1983, Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models, in: J.A. Frenkel, ed., Exchange Rates and International Macroeconomics (University of Chicago Press for the NBER, Chicago, IL).
Hodrick, Robert J. and Sanjay Srivastava 1984, An investigation of risk and return in forward foreign exchange, Journal of International Money and Finance 3, 5–29.
Hsieh, David A. 1984, Tests of rational expectations and no risk premium in forward exchange markets, Journal of International Economics 17, 173–184.
Korajczyk, Robert A. 1985, The pricing of forward contracts for foreign exchange, Journal of Political Economy 93, 346–368.
Krasker, W.S. 1980, The 'Peso Problem' in testing forward foreign exchange markets, Journal of Monetary Economics 6, 269–276.
Levi, Maurice 1978, The weekend game: Clearing house vs. Federal funds, Canadian Journal of Economics 11, 750–757.
Longworth, David, Paul Boothe and Kevin Clinton 1983, A study of the efficiency of foreign exchange markets, Oct., 92 pp., Bank of Canada.

Pagan, Adrian and Anthony D. Hall 1983, Diagnostic tests as residual analysis, Econometric Reviews 2, 159–218.

Ramsey, J.B. 1969, Tests for specification errors in classical linear least-squares regression analysis, Journal of the Royal Statistical Society, Series B, 31, 350–371.

White, Halbert 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, Econometrica 48, 817–838.

White, Halbert 1982, Maximum likelihood estimation of misspecified models, Econometrica 50, 1–25.