Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis

Lars Peter Hansen; Robert J. Hodrick

The Journal of Political Economy, Vol. 88, No. 5 (Oct., 1980), 829-853.

Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198010%2988%3A5%3C829%3AFERAOP%3E2.0.CO%3B2-J

The Journal of Political Economy is currently published by The University of Chicago Press.

Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis

Lars Peter Hansen and Robert J. Hodrick

Graduate School of Industrial Administration, Carnegie-Mellon University

This paper examines the hypothesis that the expected rate of return to speculation in the forward foreign exchange market is zero; that is, the logarithm of the forward exchange rate is the market's conditional expectation of the logarithm of the future spot rate. A new computationally tractable econometric methodology for examining restrictions on a k-step-ahead forecasting equation is employed. Using data sampled more finely than the forecast interval, we are able to reject the simple market efficiency hypothesis for exchange rates from the 1970s and the 1920s. For the modern experience, the tests are also inconsistent with several alternative hypotheses which typically characterize the relationship between spot and forward exchange rates.

I. Introduction

The purpose of this paper is twofold. First, we examine the "efficient-markets hypothesis" for foreign exchange markets for several different currencies. By the efficient-markets hypothesis, we mean the proposition that the expected rate of return to speculation in the forward foreign exchange market conditioned on available information is zero. Several authors have noted theoretical problems with this proposition since it ignores some intertemporal allocation and risk considerations. These theoretical arguments indicate that one should not equate empirical rejection of this notion of efficiency

We would like to thank Russell Boyer, Jacob Frankel, Peter Garber, John Geweke, Craig Hakkio, Rex Thompson, and members of the New York University International Economics Seminar for helpful comments. We also thank Frank Russell for assistance in performing the computations and Dale Henderson for help in securing the data.

[*Journal of Political Economy*, 1980, vol. 88, no. 5]

with evidence of market failure. However, this does not remove all interest in tests of the hypothesis. The extent to which these exchange markets can be characterized approximately as efficient markets remains an interesting question that can best be answered through formal econometric analysis. The tests we perform rely on an asymptotically more efficient estimation technique than previous researchers have used in investigating the hypothesis.

The second purpose of the paper is to implement a computationally tractable estimation procedure for examining restrictions on a k-step-ahead forecasting equation. This type of estimation problem is of particular importance since often rational expectations restrictions on a k-step-ahead forecasting equation can take on a particularly simple form, even though the implicit restrictions on the one-step-ahead equation are cumbersome. This problem arises since the sampling time interval typically is finer than the forward-contract intervals characteristic of the foreign exchange market or other organized futures markets. Direct estimation of the k-step-ahead equation is therefore desirable. Although the estimation technique we employ has no claim to full asymptotic efficiency, we indicate why, in many instances, the maximum likelihood estimators are not very computationally convenient.

The paper is divided into five sections and an Appendix. Section II contains a discussion of the efficient-markets hypothesis as applied to the foreign exchange market. Alternative estimation strategies, including the one employed in our empirical research, are investigated in Section III. Section IV presents results of the empirical investigation based on data from the modern experience with flexible exchange rates and on data from the 1920s. Section V provides interpretations of our results and outlines possible future investigations.

II. The Simple Efficiency Hypothesis

Any discussion of the efficiency of a market requires a specification of the preferences and information sets of economic agents, the technology available for production, and the costs inherent in transactions. We are interested in testing a conventional and perhaps relatively simplistic version of efficiency in the foreign exchange market. If economic agents are risk neutral, costs of transaction are zero, information is used rationally, and the market is competitive, the foreign exchange market will be efficient in the sense that the expected rate of return to speculation in the forward exchange market will be zero.

Let $s_t = \ln(S_t)$ and $f_{t,k} = \ln(F_{t,k})$, where S_t and $F_{t,k}$ are the levels of the spot exchange rate and the k-period forward exchange rate determined at time t. Since $s_{t+k} - f_{t,k}$ is an approximate measure of the rate of return to speculation, the simple efficient-markets hypothesis is that

$$f_{t,k} = E(s_{t+k}/\Phi_t) \tag{1}$$

where $E(\cdot/\Phi_t)$ signifies the mathematical expectation conditioned on the information set available to agents at time t.[1]

Because of the joint nature of the efficiency hypothesis, rejection of (1) is not immediately translatable into rejection of the efficiency or rationality of the foreign exchange market. One well-known alternative hypothesis arises if economic agents are risk averse.

Grauer, Litzenberger, and Stehle (1976) and Stockman (1978) demonstrate that risk aversion implies that, in equilibrium, the forward exchange rate equals the conditional expectation of the future spot rate plus a risk premium. Similarly, in the context of commodity markets, Danthine (1978) demonstrates conditions under which the simple efficiency hypothesis fails to hold in a dynamic competitive equilibrium in which people use all available information rationally.

Geweke and Feige (1979) state, "To be informative, an econometric procedure should be powerful enough to reject the efficient markets hypothesis, and it should provide some indication of why the hypothesis is not true in the market being studied." The econometric procedures that we advocate and employ are asymptotically more efficient than most of the traditional techniques that have been used in the study of forward exchange rates. While our procedures indicate rejection of the null hypothesis in some cases, further specification of alternative hypotheses seems necessary before we fully understand the reason for rejection. In Section V we discuss possible interpretations of our results. In Section III we examine alternative estimation strategies and discuss the power of alternative tests of the efficiency hypothesis.

III. Alternative Estimation Strategies

Here we consider the general problem of estimating the parameters of a k-step-ahead linear forecasting equation

$$E(y_{t+k}/\Phi_t) = x_t\beta \tag{2}$$

[1] Since the exchange rate is the relative price of two moneys, it can be expressed as the home currency price of a unit of foreign currency or vice versa. Note that the logarithmic expression for the rate of return to speculation and consequently eq. (1) are independent of the way exchange rates are quoted. If the hypothesis were not expressed in logarithmic form, Siegel's (1972) paradox would arise. See McCulloch (1975), Roper (1975), Boyer (1977), and Stockman (1978) for discussions of the latter issue.

where x_t is an l-dimensional row vector of variables contained in Φ_t, the information set available at time t, and β is an l-dimensional column vector of parameters. It is of interest to test whether β is equal to some hypothesized value. Variants of the forecasting-efficiency hypothesis discussed in the previous section impose restrictions on k-step-ahead forecasting equations like (2) where either the spot exchange rate or the rate of return to speculation is the variable being forecast. The exchange market efficiency tests which we discuss in Section IV are embedded in this framework. Consider the forecast error $u_{t,k} = y_{t+k} - E(y_{t+k}/\Phi_t)$. It can be verified that $E(u_{t,k}u_{t+h,k})$ is zero for all $h \geq k$. Only in the case in which the sampling interval equals the forecast interval, that is, $k = 1$, will the forecast errors be serially uncorrelated.

One way to proceed is to estimate the parameter β in the regression equation

$$y_{t+k} = x_t\beta + u_{t,k}. \tag{3}$$

It is easily verified that $E(x_t u_{t,k}) = 0$. This is the key requirement for the ordinary least squares (OLS) estimator of β to be consistent.[2] However, asymptotic justification of the conventional computation of standard errors in OLS regressions requires that the errors be serially uncorrelated. Consequently, in testing hypotheses concerning forecasting equations, one alternative is to define the sampling interval to be equal to the forecast interval. In the context of tests of exchange market efficiency, this alternative clearly does not make use of all available data. Cornell (1977), Frenkel (1977, 1978, 1979), Levich (1978), and Geweke and Feige (1979) have all used nonoverlapping samples to circumvent problems with serial correlation but have sacrificed observations in the process.

A standard econometric technique for estimation in the presence of serially correlated errors is generalized least squares (GLS). Time series versions of GLS techniques require the strict econometric exogeneity of the x_t process in (2). This means that $E(u_{t,k}/x_t,x_{t-1},x_{t+1}, \ldots) = 0$. Obviously, this condition is not implied by the fact that $u_{t,k}$ is the k-step-ahead forecast error of y_{t+k}. The strict exogeneity assumption is a claim that knowledge of future x_t's would be useless in determining the optimal forecast for y_{t+k}. In testing exchange market efficiency, researchers typically include either the contemporaneous forward exchange rate, past forecast errors, or past rates of return to speculation in the x_t vector. In each case the assumption of strict exogeneity is clearly inappropriate since knowledge of future values

[2] One possible set of requirements to ensure the consistency of the ordinary least squares estimator is: (i) y_t and x_t are jointly stationary and ergodic; (ii) $E(x_t'x_t)$ is nonsingular; and (iii) $E(x_t u_{t+k}) = 0$.

of these variables would provide useful information in forecasting future spot exchange rates, forecast errors, or rates of return. If the x_t vector contains variables which are not strictly exogenous, GLS estimation of β that implicitly filters the data distorts the orthogonality conditions and renders the estimator inconsistent.[3]

Even OLS estimation in which the regressors are endogenous presents some complications. For instance, since explicit determination of small sample properties is cumbersome if not intractable, one must rely on asymptotic distribution theory for standard error and significance level computations. Derivation of the asymptotic distribution of the parameter estimators typically exploits the assumption that the x_t and y_t processes are stationary.[4] Because time series evidence for the exchange rate processes indicates that these processes are in fact nonstationary, we have used forecast errors or rates of return in our tests of the efficiency hypothesis. Our contention is that these variables are more likely to meet the requirement of stationarity.[5]

Our strategy is to estimate β consistently with OLS procedures using data sampled more finely than the forecast interval, but we follow Hansen (1979) in making the appropriate modifications in the estimation of the asymptotic covariance matrix. In this manner we are able to increase dramatically the sample size of the data used in our tests, with corresponding gains in the asymptotic power of the tests. We note again that this estimation strategy is not fully efficient. In the Appendix we discuss more efficient estimation strategies for β which are computationally much more burdensome than the OLS estimation with modified standard errors.

The assumptions needed to justify the procedures we have used are now examined in some detail. Assume that the y_t and x_t processes have been transformed to jointly stationary and ergodic processes, and that the best linear predictor is equivalent to the conditional expectation. Let Δ_t be the information set generated by current and all past values

[3] Hansen (1979) discusses the inconsistency of GLS in this situation. Stockman (1978) has appropriately used a GLS procedure to estimate a forecasting equation with only a constant term in x_t.

[4] The standard procedure for determining small sample properties of OLS estimators by conditioning on the entire vector of right-hand-side variables is inappropriate in the absence of the strict exogeneity assumption. Fuller (1976) and Sims (1978) provide treatments of the asymptotic distribution of OLS for autoregressive specifications with some unstable roots. It also is clear that the identical distribution assumption embedded in the stationary requirement can be relaxed somewhat. Thus stationarity is not always needed to provide justification that test statistics' distributions are asymptotically normal or χ^2.

[5] Geweke and Feige (1979) use the "realized rate of exchange gain," $(S_{t+k} - F_{t,k})/S_t$, in their analysis in order to satisfy the requirement of stationarity. Frenkel (1977, 1978, 1979) regresses s_{t+k} on $f_{t,k}$, while Frenkel and Clements (1978) and Levich (1978) report regressions of S_{t+k} on $F_{t,k}$.

of y_t and x_t, and let $v_t = y_t - E(y_t/\Delta_{t-1})$ and $w_t = x_t' - E(x_t'/\Delta_{t-1})$. Note that Δ_t is a subset of Φ_t, and v_t and w_t are the one-step-ahead forecast errors for y_t and x_t using the information set Δ_{t-1}. We assume that

$$E\left[\begin{bmatrix} v_t \\ w_t \end{bmatrix}[v_t'w_t']/\Delta_{t-1}\right] = \Lambda,$$

a matrix of constants independent of the elements in Δ_{t-1}.

Under the above assumptions, Hansen (1979) demonstrates that $\sqrt{T}(\hat{\beta}_T - \beta)$ converges in distribution to a normally distributed random vector with mean zero and covariance matrix Θ where T is the sample size, $\hat{\beta}_T$ is the OLS estimator,

$$\Theta = R_x(0)^{-1}\mathfrak{S}R_x(0)^{-1},$$

$$\mathfrak{S} = \sum_{j=-k+1}^{k-1} R_u(j)R_x(j),$$

$$R_u(j) = E(u_{t,k}u_{t+j,k}),$$

and

$$R_x(j) = E(x_t'x_{t+j}).$$

In order to compute asymptotically justified confidence regions, it is necessary to obtain consistent estimators of $R_x(j)$ and $R_u(j)$ for $j = -k + 1, \ldots, k - 1$. Since x_t is ergodic for $j \geqslant 0$,

$$\hat{R}_x^T(j) = \frac{1}{T}\sum_{t=j+1}^{T} x_t'x_{t-j} \rightarrow R_x(j) \text{ almost surely.}$$

Hansen demonstrates that, for $j \geqslant 0$,

$$\hat{R}_u^T(j) = \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_{t,k}^T\hat{u}_{t-j,k}^T \rightarrow R_u(j) \text{ almost surely}$$

where $\hat{u}_{t,k}^T$ is the OLS residual for observation t with sample size T. Using the facts that $R_u(j) = R_u(-j)$ and $R_x(j) = R_x'(-j)$, we can obtain a consistent estimator of the asymptotic covariance matrix Θ.

It is useful for us to represent this proposed covariance matrix estimator in the following fashion for computational purposes. Stack the T observations on x_t into a matrix

$$X_T = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_T \end{bmatrix}$$

and form a $T \times T$ symmetric matrix $\hat{\Omega}_T$ whose lower triangular representation is

$$\begin{bmatrix} \hat{R}_u^T(0) & & & & & \\ \hat{R}_u^T(1) & \hat{R}_u^T(0) & & & & \\ \vdots & & \ddots & & & \\ \hat{R}_u^T(k-1) & & & \ddots & & \\ 0 & & & & \ddots & \\ \vdots & & & & & \ddots \\ 0 \cdots 0 & \hat{R}_u^T(k-1) & \cdots & & \hat{R}_u^T(1) & \hat{R}_u^T(0) \end{bmatrix}.$$

Noting that

$$T(X_T'X_T)^{-1} = \hat{R}_x^T(0)^{-1}$$

and

$$\frac{1}{T}(X_T'\hat{\Omega}_T X_T) = \sum_{j=-k+1}^{k-1} \hat{R}_u^T(j)\hat{R}_x^T(j),$$

we conclude that

$$T(X_T'X_T)^{-1}X_T'\hat{\Omega}_T X_T(X_T'X_T)^{-1} = \hat{\Theta}_T \tag{4}$$

is a consistent estimator for the asymptotic covariance matrix.

When the right-hand-side variables are strictly exogenous, the finite sample exact covariance matrix conditioned on X_T for $\sqrt{T}(\hat{\beta}_T - \beta)$ is

$$T(X_T'X_T)^{-1}X_T'\Omega_T X_T(X_T'X_T)^{-1} \tag{5}$$

where $\Omega_T = E(U_T U_T'/X_T)$ and

$$U_T = \begin{bmatrix} u_{1,k} \\ \cdot \\ \cdot \\ \cdot \\ u_{T,k} \end{bmatrix}.$$

Note the similarity between expressions (4) and (5). The assumptions made above provide asymptotic justification for (4) even in the absence of the strict exogeneity assumption.[6]

[6] The assumption of a constant conditional covariance matrix Λ can be relaxed. Hansen (1979) traces through the impact of relaxing this assumption in the derivation of the asymptotic covariance matrix for the OLS estimator.

For purposes of comparison, we now introduce an alternative estimator of β in (2). Suppose we sample x_t, y_t, and $u_{t,k}$ at every kth integer; that is, we form new series $x_t^* = x_{kt}$, $y_t^* = y_{kt}$, and $u_t^* = u_{kt,k}$. Note that

$$y_{t+1}^* = x_t^* \beta + u_t^*, \tag{6}$$

and that $E(u_t^* u_{t-j}^*) = 0$ for $j \neq 0$. By sampling at every kth integer, we have obtained a regression equation with a serially uncorrelated disturbance term. Under the assumptions made above, x_t^* and y_t^* are jointly stationary and ergodic, $E(u_t^*/\Phi_t^*) = 0$, and $E(u_t^{*2}/\Phi_t^*) = R_u(0)$ where $\Phi_t^* = \Phi_{kt}$. Equation (6) can be estimated using OLS and employing the sampled data (x_t^*, y_t^*) for $t = 1, 2, \ldots, T^*$. The integer T^* is the number of sampled observations available to the econometrician and satisfies $T/k \leq T^* \leq T/k + 1$. Let us denote the resulting estimator $\bar{\beta}_T$. It can be demonstrated that $\sqrt{T^*}(\bar{\beta}_T - \beta)$ converges in distribution to a normally distributed random vector with mean zero and covariance matrix Σ, where $\Sigma = R_u(0)R_x(0)^{-1}$.[7] This implies that $\sqrt{T}(\bar{\beta}_T - \beta)$ converges in distribution to a normally distributed random vector with mean zero and covariance matrix $k\Sigma$. Comparing this covariance matrix with Θ, the asymptotic covariance matrix for $\sqrt{T}(\hat{\beta}_T - \beta)$, demonstrates the sense in which there is a gain to employing our estimation strategy over OLS with nonoverlapping data. In the Appendix we demonstrate that $k\Sigma$ exceeds Θ by a positive definite matrix other than in an exceptional case in which $k\Sigma = \Theta$. Consequently, the asymptotic standard errors for $\hat{\beta}_T$ are smaller than those for $\bar{\beta}_T$.

Since we are interested in performing tests of the hypothesis that $\beta = \beta_o$ for some specified β_o, a question of considerable importance is the relative power of testing the hypothesis using $\hat{\beta}_T$ versus $\bar{\beta}_T$. Let $\hat{\Theta}_T$ and $\bar{\Sigma}_T$ be consistent estimators of the covariance matrices Θ and Σ. From the asymptotic distribution theory supplied above, we know that $T(\hat{\beta}_T - \beta_o)'\hat{\Theta}_T^{-1}(\hat{\beta}_T - \beta_o)$ and $(T/k)(\bar{\beta}_T - \beta_o)'\bar{\Sigma}_T^{-1}(\bar{\beta}_T - \beta_o)$ both approximately have χ^2 distributions with l degrees of freedom under the null hypothesis $\beta = \beta_o$. From the standpoint of relative asymptotic strength, Geweke (1979) provides formal justification for examining the probability limit of the ratio

$$\frac{T(\hat{\beta}_T - \beta_o)'\hat{\Theta}_T^{-1}(\hat{\beta}_T - \beta_o)}{(T/k)(\bar{\beta}_T - \beta_o)'\bar{\Sigma}_T^{-1}(\bar{\beta}_T - \beta_o)} \tag{7}$$

in circumstances in which $\beta \neq \beta_o$.[8] This ratio converges almost surely to

[7] The derivation of the asymptotic distribution of $\bar{\beta}_T$ can be viewed as a special case of Hansen (1979) and is the standard result of OLS estimation with serially uncorrelated disturbances.

[8] See Geweke 1979, theorem, p. 9.

$$\frac{(\beta - \beta_o)'\Theta^{-1}(\beta - \beta_o)}{(1/k)(\beta - \beta_o)'\Sigma^{-1}(\beta - \beta_o)} \tag{8}$$

under the assumptions specified for (2). Under this range of alternative hypotheses, we demonstrate in the Appendix that (8) is greater than or equal to one, allowing us to assert that the procedure which employs the complete data set dominates the sampling procedure. This raises the issue of whether under other alternative hypotheses the sampling procedure dominates the one employing all of the data.

Suppose we maintain the assumptions that y_t and x_t have finite second moments and are jointly stationary and ergodic, but we eliminate the assumption that $E(y_{t+k}/\Phi_t) = x_t\beta$. Instead, define β to be the parameter vector such that $x_t\beta$ is the linear least squares projection of y_{t+k} on x_t.[9] We still are assured that $E(x_t u_{t,k}) = 0$ in (3) and that $E(x_t^* u_t^*) = 0$ in (6). These guarantee that $\hat{\beta}_T$ and $\bar{\beta}_T$ converge almost surely to β, although it now will not necessarily be true that $E(u_{t,k}u_{t+h,k}) = 0$ for $h \geqslant k$ or that $E(u_t^* u_{t+h}^*) = 0$ for all $h \neq 0$. With the loss of these implications the matrix $k\Sigma$ may be less than Θ; thus the ratio (8) may be less than one. In these circumstances $k\Sigma$ and Θ are not asymptotically justifiable covariance matrices for $\bar{\beta}_T$ and $\hat{\beta}_T$, respectively. Since there are alternative hypotheses under which $k\Sigma$ may be less than Θ, we computed estimates of both matrices in our empirical analysis.[10]

IV. Empirical Results

We conducted tests of the efficiency of the foreign exchange market for the modern experience with flexible exchange rates which began essentially in March 1973 and for the period of generalized floating exchange rates following World War I. In the first part of this section we present our results for the modern period.

Data were obtained for the spot and 3-month forward exchange rates for seven currencies: the Canadian dollar, the deutsche mark, the French franc, the U.K. pound, the Swiss franc, the Japanese yen, and the Italian lira. All exchange rates are expressed in U.S. cents per

[9] There are two major distinctions between the expectation of y_{t+k} conditioned on Φ_t and the linear least squares projection of y_{t+k} on x_t. First, the information set Φ_t contains much more than just x_t; i.e., past values of x_t and current and past values of y_t are also included. Second, the conditional expectation need not equal a linear combination of the elements of Φ_t.

[10] It is straightforward to construct a test which dominates the OLS procedure with nonoverlapping data under a broader range of alternatives and is a minor modification of our test. Let $S_T = \max[T(\hat{\beta}_T - \beta_o)'\hat{\Theta}_T^{-1}(\hat{\beta}_T - \beta_o), (T/k)(\bar{\beta}_T - \beta_o)'\bar{\Sigma}_T^{-1}(\bar{\beta}_T - \beta_o)]$. Under the null hypothesis, S_T is asymptotically χ^2 distributed with l degrees of freedom. It can be shown that $S_T/(T/k)(\bar{\beta}_T - \beta_o)'\bar{\Sigma}_T^{-1}(\bar{\beta}_T - \beta_o)$ converges almost surely to a number greater than or equal to one under alternative hypotheses where $\beta \neq \beta_o$ as long as y_t and x_t have finite second moments and are stationary and ergodic.

unit of foreign currency. The data were sampled to form a weekly series, as described in the Data section of the Appendix.[11]

The hypothesis that $f_{t,k} = E(s_{t+k}/\Phi_t)$ implies that the forecast error $s_{t+k} - f_{t,k}$ is uncorrelated with information available at time t. We tested this hypothesis using ordinary least squares regression with the modified standard errors as described in Section III. Although any element of the information set, Φ_t, could be used in a test of the hypothesis that $s_{t+k} - f_{t,k}$ is orthogonal to Φ_t, in order to have a powerful test one would want to use elements which are a priori likely to be important under alternative hypotheses. The elements of Φ_t which we have chosen are the most recent past forecast errors from the own exchange rate and other exchange rates. We employ two forms of alternative specifications against which we test the null hypothesis. Using terminology from the efficient-markets literature, we can classify the first as a test of the weak form of the efficient-markets hypothesis since only the past forecast errors from the own exchange rate are allowed to have nonzero coefficients. The second form is labeled a semistrong test since past forecast errors from other exchange markets are included in the model. This second specification closely parallels the hypothesis-testing strategy employed by Geweke and Feige (1979).[12]

In table 1 we present the estimated regressions of the forecast error on a constant and two lagged forecast errors, using weekly data and a 3-month or 13-week forward rate, as in

$$s^i_{t+13} - f^i_t = a_i + b_{i1}(s^i_t - f^i_{t-13}) + b_{i2}(s^i_{t-1} - f^i_{t-14}) + u^i_t \tag{9}$$

for $i = 1, \ldots, 7$ currencies. We test the joint hypotheses that a_i, b_{i1}, and b_{i2} are all zero. The most damaging evidence against the null hypothesis occurs in the deutsche mark–U.S. dollar exchange rate. Although the constant is insignificantly different from zero, the two lagged forecast errors have marginal significance levels smaller than .02. For the Swiss franc and the Italian lira we find individual coefficient estimates that have marginal significance levels less than .1. The other currencies provide no strong evidence against the null hypothesis.

Table 2 gives results of the regression of the forecast error for a

[11] The data were obtained from the Board of Governors of the Federal Reserve System. See the Data section of the Appendix for a complete description of the series used in this study.

[12] Geweke and Feige (1979) have used this specification with the realized rate of exchange gain as defined in n. 3 and employing nonoverlapping data but testing several markets simultaneously. Our procedures could be modified to allow multiequation tests.

TABLE 1

$$(s^i_{t+13} - f^i_t) = a_i + b_{i1}(s^i_t - f^i_{t-13}) + b_{i2}(s^i_{t-1} - f^i_{t-14}) + u^i_t$$

Currency	Sample Period	N Observations	$\hat{a}_i$ Confidence	$\hat{b}_{i1}$ Confidence	$\hat{b}_{i2}$ Confidence	χ^2 (3) All Coefficients = 0 Confidence	R^2	Residual Variance
1. Canadian dollar	10/9/73–1/16/79	276	−.004 (.004) .610	.059 (.213) .217	.226 (.213) .710	4.065 .745	.076	.0004
2. Deutsche mark	10/9/73–1/16/79	276	.008 (.011) .560	.662 (.269) .986*	−.910 (.272) .999*	11.966 .992*	.135	.0027
3. French franc	10/9/73–1/16/79	276	.007 (.011) .472	.269 (.220) .778	−.072 (.218) .258	2.609 .544	.043	.0026
4. U.K. pound	10/9/73–1/16/79	276	.001 (.011) .084	.385 (.317) .776	−.423 (.313) .823	1.860 .398	.015	.0026
5. Swiss franc	10/9/73–1/16/79	276	.022 (.015) .851	.294 (.230) .798	−.459 (.227) .957*	5.595 .867	.040	.0045
6. Japanese yen	7/9/74–1/16/79	236	−.023 (.046) .391	.004 (.048) .060	−.046 (.041) .734	4.277 .767	.031	.0021
7. Italian lira	4/8/75–1/16/79	197	.008 (.017) .380	.690 (.406) .911*	−.722 (.407) .924*	3.510 .680	.037	.0040

NOTE.—Figures in parentheses = SE. Interpretation of these values as SE requires the assumption that the regression equation contains the correct expression for the conditional expectation of the left-hand-side variable conditioned on Φ_t. Confidence is 1 minus the marginal significance level. Values of the confidence term which are close to 1 indicate evidence against the null hypothesis.
* Marginal significance level < .1.

TABLE 2

$$(s^i_{t+13} - f^i_t) = a_i + \sum_{j=1}^{5} b_{ij}(s^j_t - f^j_{t-13}) + u^i_t$$

Currency	Sample Period	N Observations	$\hat{a}_i$ Confi-dence	$\hat{b}_{i1}$ Confi-dence	$\hat{b}_{i2}$ Confi-dence	$\hat{b}_{i3}$ Confi-dence	$\hat{b}_{i4}$ Confi-dence	$\hat{b}_{i5}$ Confi-dence	χ^2 (6) All Coeffi-cients = 0 Confi-dence	R^2	Residual Variance
1. Canadian dollar	10/2/73–1/16/79	277	−.004 (.004) .731	.033 (.145) .182	−.045 (.088) .391	−.161 (.081) .954*	−.095 (.070) .823	.051 (.077) .496	20.508 .998*	.334	.0003
2. Deutsche mark	10/2/73–1/16/79	277	.003 (.011) .201	.147 (.422) .273	−.665 (.260) .989*	−.224 (.237) .656	.201 (.203) .677	.563 (.226) .987*	12.390 .946*	.276	.0023
3. French franc	10/2/73–1/16/79	277	.003 (.011) .194	−.067 (.417) .127	−.358 (.256) .838	.055 (.234) .185	.174 (.201) .613	.401 (.223) .928*	7.561 .728	.176	.0023
4. U.K. pound	10/2/73–1/16/79	277	−.004 (.012) .269	−.650 (.453) .848	−.100 (.281) .277	−.269 (.255) .708	−.087 (.219) .310	.255 (.244) .704	3.882 .307	.085	.0025
5. Swiss franc	10/2/73–1/16/79	277	.020 (.013) .867	.628 (.526) .768	−.977 (.324) .997*	.237 (.295) .578	.225 (.253) .627	.362 (.282) .801	13.872 .969*	.256	.0036

NOTE.—See table 1. The coefficient b_{ij} refers to the regression coefficient of currency i on currency j where the i and j subscripts index the currencies as shown above.
* Marginal significance level < .1.

currency on lagged values of the own forecast error and four other currencies' lagged forecast errors, as in

$$s^i_{t+13} - f^i_t = a_i + \sum_{j=1}^{5} b_{ij}(s^j_t - f^j_{t-13}) + u^i_t \tag{10}$$

for $i = 1, \ldots, 5$ currencies. The null hypothesis that all coefficients in the regression are zero is rejected for the Canadian dollar, the deutsche mark, and the Swiss franc at all significance levels greater than .06. Since the dependent variable in table 2 covers the same sample period as that in table 1, the multicountry test appears to be a more powerful test of the efficiency hypothesis because we are able to reject the hypothesis for two other countries except at very low significance levels. The results in table 2 indicate that lagged forecast errors for some currencies have explanatory power in predicting the current forecast errors for the three currencies mentioned above.

The analysis in tables 1 and 2 uses data from the earliest part of the modern flexible exchange rate experience which includes the especially volatile interval from October 1973 to February 1974 (the period of the Arab-Israeli War and the oil boycott). Since learning in the market about the functioning of the new flexible exchange rate system and the disruption of the war and the oil crisis might have biased our previous tests, we reestimated equation (10) for a period running from June 25, 1974, to January 16, 1979. The results in table 3 for the shorter sample could be more representative of the situation under the mature float of today. In the joint test of all six coefficients, only the Canadian dollar has a marginal significance level below .1, and it is below .01. In the deutsche mark, the U.K. pound, and the Swiss franc regressions, there are individual coefficients with marginal significance levels less than .1. If the alternative hypothesis were true, some increase in the marginal significance level would be expected as the sample size decreased. The increase in the case of the deutsche mark appears to be more than can be accounted for by this explanation, suggesting evidence in favor of the null hypothesis.[13]

In table 4 we expand the multicountry test to include the Japanese yen and the Italian lira, for which data are available only since April 1, 1975. This inclusion has a dramatic effect on the test of the efficiency hypothesis. We again reject the hypothesis for the Canadian dollar except at extremely low significance levels. However, the marginal significance level for the deutsche mark has fallen to .001, and it is .09

[13] Leamer (1978) argues persuasively that, from a Bayesian point of view, the critical significance level for rejection of the null hypothesis should be a decreasing function of sample size given that one assigns positive a priori probability that the null hypothesis is true.

TABLE 3

$$(s^i_{t+13} - f^i_t) = a_i + \sum_{j=1}^{5} b_{ij}(s^j_t - f^j_{t-13}) + u^i_t$$

Currency	Sample Period	N Observations	$\hat{a}_i$ Confidence	$\hat{b}_{i1}$ Confidence	$\hat{b}_{i2}$ Confidence	$\hat{b}_{i3}$ Confidence	$\hat{b}_{i4}$ Confidence	$\hat{b}_{i5}$ Confidence	χ^2 (6) All Coefficients = 0 Confidence	R^2	Residual Variance
1. Canadian dollar	6/25/74–1/16/79	238	−.005 (.004) .788	.034 (.152) .177	−.126 (.112) .741	−.188 (.097) .948*	−.072 (.075) .661	.100 (.086) .754	18.030 .994*	.298	.0003
2. Deutsche mark	6/25/74–1/16/79	238	.008 (.012) .515	.035 (.399) .070	−.510 (.305) .905*	−.168 (.263) .476	.176 (.192) .641	.357 (.238) .866	5.169 .478	.135	.0018
3. French franc	6/25/74–1/16/79	238	.011 (.010) .742	−.128 (.348) .287	−.037 (.262) .113	.158 (.227) .516	.120 (.169) .523	.101 (.203) .382	9.036 .828	.147	.0015
4. U.K. pound	6/25/74–1/16/79	238	−.007 (.013) .394	−.760 (.445) .912*	.231 (.337) .506	.003 (.291) .008	−.225 (.218) .699	.023 (.261) .071	4.191 .349	.109	.0022
5. Swiss franc	6/25/74–1/16/79	238	.025 (.016) .879	.598 (.554) .719	−.868 (.421) .961*	.319 (.363) .620	.227 (.270) .601	.189 (.327) .438	8.591 .802	.162	.0036

NOTE.—See tables 1 and 2.
* Marginal significance level < .1.

TABLE 4

$$(s^i_{t+13} - f^i_t) = a_i + \sum_{j=1}^{7} b_{ij}(s^j_t - f^j_{t-13}) + u^i_t$$

Currency	Sample Period	N Obser-vations	$\hat{a}_i$ Confi-dence	$\hat{b}_{i1}$ Confi-dence	$\hat{b}_{i2}$ Confi-dence	$\hat{b}_{i3}$ Confi-dence	$\hat{b}_{i4}$ Confi-dence	$\hat{b}_{i5}$ Confi-dence	$\hat{b}_{i6}$ Confi-dence	$\hat{b}_{i7}$ Confi-dence	χ^2 (8) All Coeffi-cients = 0 Confi-dence	R^2	Residual Variance
1. Canadian	4/1/75–		−.001	−.103	−.228	−.117	−.018	.154	−.164	−.076	26.085		
dollar	1/16/79	198	(.004)	(.164)	(.121)	(.102)	(.084)	(.086)	(.093)	(.064)	.999*	.402	.0003
			.237	.469	.942*	.747	.172	.928*	.920*	.765			
2. Deutsche	4/1/75–		−.001	.671	−.029	−.500	.145	−.003	.631	.247	36.688		
mark	1/16/79	198	(.007)	(.269)	(.210)	(.176)	(.136)	(.146)	(.154)	(.107)	.999*	.508	.0009
			.098	.987*	.108	.995*	.713	.018	.999*	.979*			
3. French	4/1/75–		.002	.052	.267	−.059	.116	−.134	.280	.809	11.990		
franc	1/16/79	198	(.008)	(.309)	(.237)	(.200)	(.158)	(.166)	(.177)	(.122)	.848	.249	.0011
			.222	.133	.740	.232	.536	.582	.886	.491			
4. U.K.	4/1/75–		−.020	−.683	.362	−.183	−.383	−.183	.589	.098	9.411		
pound	1/16/79	198	(.014)	(.470)	(.383)	(.320)	(.240)	(.263)	(.273)	(.192)	.691	.235	.0022
			.856	.854	.656	.433	.889	.513	.969*	.390			
5. Swiss	4/1/75–		.007	.847	−.445	.102	.155	−.217	.799	−.014	13.793		
franc	1/16/79	198	(.017)	(.541)	(.461)	(.383)	(.275)	(.313)	(.316)	(.225)	.913*	.292	.0030
			.324	.882	.666	.210	.426	.513	.988*	.051			
6. Japanese	4/1/75–		.023	.225	.108	.262	−.073	−.217	.112	.006	6.416		
yen	1/16/79	198	(.014)	(.451)	(.388)	(.322)	(.229)	(.263)	(.265)	(.188)	.399	.065	.0020
			.897	.381	.220	.585	.250	.592	.327	.025			
7. Italian	4/1/75–		−.011	−.740	.874	−.302	−.326	−.421	.745	−.118	12.476		
lira	1/16/79	198	(.015)	(.545)	(.431)	(.361)	(.279)	(.299)	(.314)	(.219)		.272	.0031
			.527	.826	.957*	.597	.756	.841	.982*	.409	.869		

NOTE.—See tables 1 and 2.
* Marginal significance level < .1.

for the Swiss franc. The coefficient for the lagged forecast error of the Japanese yen is significantly different from zero in four countries at the .04 level and in the fifth at the .08 level. Failure of the efficiency hypothesis is also manifest in relative high values of the R^2 statistic for the Canadian dollar and the deutsche mark.[14]

We turn now to an investigation of the other major historical experience with flexible exchange rates which occurred after World War I. Table 5 presents evidence of our tests of the efficiency hypothesis for three currencies relative to the U.K. pound. The data are weekly spot and 1-month forward rates and were obtained from Einzig (1937). This period coincides with the well-studied German hyperinflation. Our results in a regression like (9) where the forecast interval is now four rather than 13 indicate that the constant term is significantly different from zero at the .01 level for the entire sample. Since the mark was experiencing a very rapid depreciation in the final weeks of the sample, we were concerned that nonstationarity in the forecast errors could invalidate our tests. We truncated the sample by the final 15 weeks and reestimated the equation. The constant term remained significantly different from zero at the .01 level. The decrease in residual variance for the shorter sample was almost twofold, indicating that our concern with nonstationarity was justified.

The experience of the French franc following World War I presents a relatively long period of floating rates. We tested the efficiency hypothesis with a regression similar to the mark analysis above. For the entire sample, January 1922 to July 1926, we reject the null hypothesis except at significance levels less than .02, finding that the constant term and the coefficients on the two lagged forecast errors are not all zero. The analysis of the estimated errors indicated that one observation yielded a 7-standard-deviation residual. This observation corresponded to the experience, commonly known as "Poincaré's bear squeeze," which occurred in March 1924.

[14] Since the number of 13-week nonoverlapping intervals in 198 weekly observations is approximately 15, we realize that the number of degrees of freedom in the calculations of the standard errors may be smaller than that which is sufficient to justify the application of asymptotic distribution theory. Our calculations indicate that using the weekly rather than quarterly data in table 4 has reduced the variances of the coefficients by approximately a factor of 2. Consequently, the equivalent nonoverlapping quarterly sample size necessary for this reduction is approximately 30. Since the standard degrees-of-freedom correction is $T/T - l$, for our purposes this would amount to inflating the standard errors by $\sqrt{30/22} \doteq 1.17$. Even in the case of nonoverlapping data, due to the endogeneity of the regressors, the $T/T - l$ degrees-of-freedom correction cannot be formally justified from the standpoint of small sample properties. The low levels of marginal significance for the Canadian dollar and the deutsche mark will be essentially unchanged by this heuristic degrees-of-freedom correction. Obviously, the reduction in the standard error in employing weekly rather than quarterly data varies across coefficients. See n. 16 for a discussion of this issue.

TABLE 5

$$s^i_{t+4} - f^i_t = a_i + b_{i1}(s^i_t - f^i_{t-4}) + b_{i2}(s^i_{t-1} - f^i_{t-5}) + u^i_t$$

Currency per U.K. Pound	Sample Period	N Observations	$\hat{a}_i$ Confidence	$\hat{b}_{i1}$ Confidence	$\hat{b}_{i2}$ Confidence	χ^2 (3) All Coefficients = 0 Confidence	R^2	Residual Variance
1. German mark	1/21/22–9/1/23	85	.343 (.137) .988*	.419 (.268) .882	−.220 (.271) .583	15.196 .998*	.051	29.194
2. German mark	1/21/22–3/24/23	70	.320 (.105) .998*	.098 (.221) .343	−.344 (.220) .882	10.822 .987*	.080	16.975
3. French franc	1/21/22–7/31/26	237	.019 (.009) .963*	.233 (.122) .945*	−.232 (.122) .942*	9.402 .976*	.022	.0049
4. French franc	1/21/22–2/23/24	110	.020 (.010) .952*	.311 (.178) .920*	−.253 (.180) .840	8.168 .957*	.035	.0027
5. French franc	6/28/24–7/31/26	110	.024 (.009) .992*	−.094 (.126) .544	.180 (.129) .835	13.245 .996*	.020	.0018
6. U.S. dollar	1/21/22–4/25/26	171	.002 (.002) .653	.141 (.099) .848	−.016 (.095) .137	3.252 .646	.020	.0002

NOTE.—See table 1.
* Marginal significance level < .1.

After World War I the French economy was in disarray.[15] The very large public debt which had been incurred to finance the war implied that the government had either to increase taxation or to resort to inflationary finance to balance its budget. The matter was complicated by the failure of Germany to make its reparations payments and by scandals in the French Treasury and the Bank of France. During 1922 and 1923 the franc depreciated relative to the pound from 51.47 in January 1922 to 84.82 on December 29, 1923. By March 8, 1924, the franc stood at 117.00. This rapid depreciation caused Poincaré to negotiate secret loans in Britain and the United States. These loans were used to buy massive quantities of francs, driving the exchange rate to 81.25 by March 22, with a further appreciation to 67.87 on May 3. After this period the franc more or less depreciated steadily, reaching a peak of 201.25 on July 17, 1926. The results of estimating two regressions for the periods prior to and after the occurrence of the bear squeeze are also presented in table 5. We can reject the efficiency hypothesis in the first case at levels of significance greater than .05 and for the second subsample at levels greater than .01. In each case the estimated constant term is positive, and its magnitude casts doubt on the null hypothesis. This indicates, as in the case of the German mark, that the logarithm of the forward exchange rate was an underestimate of the logarithm of the subsequently observed spot rate. Analysis of the U.S. dollar–U.K. pound exchange rates during the early 1920s provides no evidence against the null hypothesis.

In the next section we provide possible interpretations of the results which we have found regarding the simple efficiency hypothesis.

V. Interpretations and Conclusions

The analysis conducted in the previous section indicates that the simple efficiency hypothesis is suspect for several currencies for the modern experience with flexible exchange rates and for the experience in the 1920s. The tests employed a new and asymptotically more powerful estimation technique than ordinary least squares with nonoverlapping data.[16] This section examines possible reasons for the

[15] Dulles (1929) provides an excellent account of the French economy during this period.

[16] The discussions in Section III and in n. 14 indicate that it is interesting to compare the relative gain in using the variances of coefficients computed in $\hat{\Theta}_T$ and the variances computed in $k\bar{\Sigma}_T$. The maximum and minimum ratios of the diagonal elements of $\hat{\Theta}_T$ to $k\bar{\Sigma}_T$ for the constant terms and the other coefficients are presented in the table below. These computations are obviously subject to estimation error, and their justification depends on the auxiliary assumptions we made in determining the form of the asymptotic covariance matrix. The ratios for table 1 are suspiciously high, but this may merely

results and suggests alternative empirical modeling strategies which could resolve remaining questions.

First, we recognize that Grauer et al. (1976) indicate that, from the standpoint of an international Sharpe-Lintner capital asset pricing model, the forward exchange rate equals the expected future spot rate plus a term arising from a possible covariance between the future exchange rate and real gross world product adjusted for risk and scaled by the world commodity price deflator. To the extent that this covariance is stable during the sample period, one might expect a significant constant term to emerge in our regressions. This provides a potential explanation for the constant terms with low marginal levels of significance which we found for the German mark and the French franc in the 1920s. Partial equilibrium analyses of the relationship between the forward exchange rate and the expected future spot rate often obtain the result that a risk premium will separate these two rates. This risk premium provides a positive expected return to those who buy a more risky currency forward. Under such a partial equilibrium analysis, the constant terms in our regressions should be negative if the German mark and the French franc were considered more risky than the U.K. pound during the 1920s. The forward rate of marks/pound or francs/pound would be above the future spot rate, implying that one could buy marks or francs forward expecting a profit. Since the signs of the constant terms are positive, this partial equilibrium analysis can account for the significance of the constant terms only under the extremely unlikely case that the pound was considered the more risky currency. In our analysis of the modern experience, we find no constant terms with low marginal levels of significance. Hence, alternative explanations must be examined. Fluctuations in risk premiums or the covariance term mentioned above could account for the vector autocorrelations found in the modern period.

---

be evidence against the null hypothesis since some of the auxiliary assumptions are implied by the null hypothesis.

Table	Constant Max	Constant Min	Other Coefficients Max	Other Coefficients Min
1	1.82	.96	11.76	5.45
2	1.38	1.09	2.14	1.35
3	1.21	1.05	2.26	1.28
4	2.15	1.22	2.47	1.33
5	1.10	1.00	3.87	2.67

Roll and Solnik (1977) have estimated a "pure foreign exchange" asset-pricing model. The index-model approach which they use explains contemporaneous correlation between forecast errors of different currencies because each forecast error is correlated with a weighted average of the forecast errors of all currencies. Roll and Solnik implement their theory by constructing a weighted average of observed forecast errors and using this as a proxy for the true measure. An alternative approach would be to estimate a latent variable model in which the expected return from holding a portfolio of all currencies is considered as an unobservable variable which influences the return to speculation in all currencies simultaneously. Finally, movement of the index through time also could be a possible explanation for the observed multivariate autocorrelation which we find in the modern period.

There are two other quite different explanations of the rejection of the simple efficiency hypothesis for forward foreign exchange markets, during both the 1920s and the modern period. The first one is that we use large sample approximations in computing probabilities associated with our test statistics without explicit knowledge of how large the sample size must be before the approximations become good. This problem is not unique to our work since it plagues much of time series analysis. The second explanation involves the implicit need to specify what economic agents know about the stochastic properties underlying important policy actions, such as how monetary policy is conducted, and when or whether capital controls will be applied.[17] Even though economic agents may process information optimally, the correct stochastic specification of government actions may not be consistent with the statistical model underlying our test. For instance, if economic agents correctly perceive that governmental actions will be roughly constant over relatively long periods and yet may change dramatically either at uncertain points in time or by an uncertain magnitude, we conjecture that it is possible that the statistical procedure we employ might yield sample autocorrelations in forecast errors that are large relative to their estimated standard errors even if the simple efficiency hypothesis is true. The cause of this could be a combination of incorrect assumptions we have made in determining the asymptotic covariance matrix for our estimators, a small sample size relative to the movements in government policy variables, and the inappropriateness of an ergodicity assumption in an environment

[17] Kareken and Wallace (1978) and Nickelsburg (1979) demonstrate the potential importance of capital controls or the anticipation of future controls in the determination of exchange rates.

where agents may assign positive a priori probabilities to events that may ultimately never occur.[18]

We indicated in the Introduction that rejection of the hypothesis cannot be identified readily with inefficiency in the exchange market. At a theoretical level, what seems to be needed is a multicountry dynamic competitive equilibrium model, perhaps built along the lines of Danthine (1978) or Lucas (1978), richly enough specified to indicate the salient influences which can cause divergence from the simple efficiency hypothesis for exchange rates.

Since the exchange rate is the relative price of two moneys, the model should have a nontrivial role for these moneys in the economies being studied. Such a model would serve as a vehicle for interpretation of empirical research on exchange markets.

Finally, in future research we plan to examine the restrictions on the joint process consisting of the spot rate and the 1- and 3-month forward rates. We also can perform tests similar to those of this paper, using the 1-month forward rate forecast errors. Since the forecast error variance for the 1-month forward rates is smaller than for 3-month forward rates, tests using the 1-month rates should be more revealing. While we have found evidence indicating that the forward rate is not the optimal predictor of the future rate for some currencies, we have not necessarily isolated time-invariant forecasting rules which improve upon the predictions of the forward rate. The 1-month forward rate data will facilitate the search for time-invariant forecasting rules, allow examination of the stability of the multicountry autocorrelations, and provide a better sample in which the issues regarding the stability of government policy and efficiency of the foreign exchange market can be addressed.

Appendix

In this Appendix we discuss in more detail the alternative estimation strategies proposed in Section III above. First, we indicate the computational complexity of the efficient maximum likelihood estimation of β.

Consider the regression equation (3) reproduced here as

$$y_{t+k} = x_t\beta + u_{t,k}, \tag{A1}$$

[18] Demonstrating the precise sense in which these conjectures are correct is beyond the scope of the present paper. In this regard, though, Frenkel (1977) has argued that inexperience of economic agents with an environment of hyperinflation may be an appropriate explanation of the forward premium's underprediction of the rate of inflation in Germany during the 1920s. Flood and Garber (1980) have noted the importance of anticipations of monetary reform in an explicit Bayesian model of the German hyperinflation.

where $u_{t,k}$ satisfies the condition $E(u_{t,k}/\Phi_t) = 0$, Φ_t contains the information set $(x_t, x_{t-1}, \ldots, y_t, y_{t-1}, \ldots)$, and x_t is an l-dimensional row vector. Maximum likelihood estimation of β requires that we trace through the joint covariance properties for the (y_t, x_t) processes which we assume to be multivariate normal. Consider the Wold decomposition of (y_t, x_t) under the assumption that the joint process is covariance stationary.[19] Then

$$y_t = \sum_{i=0}^{\infty} \delta_i v_{t-i} + \sum_{i=0}^{\infty} \gamma_i w_{t-i} + d_{yt}, \tag{A2}$$

$$x_t' = \sum_{i=0}^{\infty} \alpha_i v_{t-i} + \sum_{i=0}^{\infty} \psi_i w_{t-i} + d_{xt}, \tag{A3}$$

where $v_t = y_t - E(y_t/y_{t-1}, y_{t-2}, \ldots, x_{t-1}, x_{t-2}, \ldots)$, $w_t = x_t' - E(x_t'/y_{t-1}, y_{t-2}, \ldots, x_{t-1}, x_{t-2}, \ldots)$, $\delta_o = 1$, $\gamma_o = 0$, $\alpha_o = 0$, $\psi_o = I$, the l-dimensional identity matrix, and d_{yt} and d_{xt} are the deterministic components of y_t and x_t, respectively. Using (A2), the conditional prediction of y_{t+k} is

$$E(y_{t+k}/y_t, y_{t-1}, \ldots, x_t, x_{t-1}, \ldots) = \sum_{i=k}^{\infty} \delta_i v_{t+k-i} + \sum_{i=k}^{\infty} \gamma_i w_{t+k-i} + d_{yt+k}. \tag{A4}$$

Using (A3), the process $x_t\beta$ is

$$x_t\beta = \sum_{i=0}^{\infty} \beta'\alpha_i v_{t-i} + \sum_{i=0}^{\infty} \beta'\psi_i w_{t-i} + \beta' d_{xt}. \tag{A5}$$

Since $E(y_{t+k}/y_t, y_{t-1}, \ldots, x_t, x_{t-1}, \ldots) = x_t\beta$, we obtain the cross-equation restrictions

$$\beta'\alpha_i = \delta_{i+k},$$
$$\beta'\psi_i = \gamma_{i+k}, \qquad i = 0, 1, 2, \ldots \tag{A6}$$

and

$$d_{yt+k} = \beta' d_{xt}.$$

In the case in which y_t and other lagged y's enter the x_t vector, additional cross-equation restrictions arise.

Assuming knowledge of d_{xt}, the joint covariance properties of the series are totally characterized by the δ_i's, the γ_i's, the α_i's, the ψ_i's, and the parameters of the covariance matrix of v_t and w_t. Maximum likelihood estimation requires the representation of these parameters subject to the restrictions (A6) in terms of a finite vector of parameters ζ which must be estimated concurrently with β. The likelihood function must be maximized via some iterative technique. The estimation of large vector autoregressive moving average processes subject to cross-equation restrictions is computationally much less tractable than the OLS procedure with modified standard errors which we use.

It was asserted in the text that $k\Sigma$, the covariance matrix for OLS with sampled nonoverlapping data, exceeds Θ, the covariance matrix for OLS with overlapping data, by a positive semidefinite matrix. We now demonstrate this proposition.

Let $q_t = u_{t,k}x_t'$, which is assumed to be stationary and to have a finite second

[19] See Sargent (1979) for a discussion of Wold's decomposition theorem and its application to time series econometrics. Rozanov (1967) provides a rigorous mathematical treatment of the theorem.

moment. Under the hypothesis that $E(u_{t,k}/\Phi_t) = 0$, it follows that $E(q_t q'_{t+j}) = 0$ for $j \geq k$. Thus we can represent q_t as a $(k - 1)$ order moving average of a vector white-noise process with covariance matrix equal to the identity matrix. In particular, write

$$q_t = \sum_{i=0}^{k-1} \eta_i e_{t-1}$$

where $E(e_t) = 0$, $E(e_t e'_t) = I$, and $E(e_t e'_{t+j}) = 0$. Define

$$\eta(z) = \sum_{i=0}^{k-1} \eta_i z^i.$$

The covariance generating function for q_t is $\eta(z)\eta(z^{-1})'$. It can be verified that

$$\eta(z)\eta(z^{-1})' = \sum_{j=1}^{k-1} \sum_{i=0}^{j-1} (\eta_i + \eta_j z^{j-i})(\eta_i + \eta_j z^{i-j})' - (k - 2)\left(\sum_{j=0}^{k-1} \eta_j \eta'_j\right). \quad \text{(A7)}$$

Thus

$$\begin{aligned} k\left(\sum_{j=0}^{k-1} \eta_j \eta'_j\right) - \eta(z)\eta(z^{-1})' &= (2k - 2)\left(\sum_{j=0}^{k-1} \eta_j \eta'_j\right) \\ &\quad - \sum_{j=1}^{k-1} \sum_{i=0}^{j-1} (\eta_i + \eta_j z^{j-i})(\eta_i + \eta_j z^{i-j})' \qquad \text{(A8)} \\ &= \sum_{j=1}^{k-1} \sum_{i=0}^{j-1} (\eta_i - \eta_j z^{j-i})(\eta_i - \eta_j z^{i-j})'. \end{aligned}$$

The first term on the left-hand side of (A8),

$$k\left(\sum_{j=0}^{k-1} \eta'_j \eta_j\right),$$

$= kE(q_t q'_t) = kR_u(0)R_x(0)$ and $\eta(z)\eta(z^{-1})'$ evaluated at $z = 1$ is the matrix $\mathfrak{S}$ in the covariance matrix Θ. Thus

$$kR_u(0)R_x(0) - \mathfrak{S} = \sum_{j=1}^{k-1} \sum_{i=0}^{j-1} (\eta_i - \eta_j)(\eta_i - \eta_j)', \quad \text{(A9)}$$

which is positive definite unless $\eta_i = \eta_j$ for all i and j between 0 and $k - 1$. Premultiplication and postmultiplication of both sides of (A9) by the symmetric matrix $R_x(0)^{-1}$ proves that $k\Sigma - \Theta$ is positive definite other than in the exceptional case when $\eta_i = \eta_j$ for all i and j between 0 and $k - 1$.

It also was asserted in the text that the ratio of the two quadratic forms, $(\beta - \beta_o)'\Theta^{-1}(\beta - \beta_o)$ and $(1/k)(\beta - \beta_o)'\Sigma^{-1}(\beta - \beta_o)$, is greater than or equal to one. This follows immediately from the fact that $\Theta^{-1} - (1/k)\Sigma^{-1}$ is positive semidefinite.

Data

The data for the modern experience with flexible exchange rates were obtained from the Board of Governors of the Federal Reserve System. Daily observations for the spot exchange rate and the 3-month forward premium expressed in percent at an annual rate were available. The premium was converted into a 3-month forward exchange rate by dividing the premium by

400, adding one, and multiplying by the spot rate since the premium was calculated as $400(F_{t,k} - S_t/S_t)$. A weekly series was constructed by taking the observation on Tuesday of each week. If no Tuesday observation was available, we used the Wednesday observation.

The data from the experience with flexible exchange rates following World War I were taken from Einzig (1937). Weekly observations on the spot exchange rates and the 1-month forward rates are available after November 19, 1921.

References

Boyer, Russell S. "The Relation between the Forward Exchange Rate and the Expected Future Spot Rate." *Intermountain Econ. Rev.* 8 (Spring 1977): 14–21.

Cornell, Bradford. "Spot Rates, Forward Rates and Exchange Market Efficiency." *J. Financial Econ.* 5 (August 1977): 55–65.

Danthine, Jean-Pierre. "Information, Futures Prices, and Stabilizing Speculation." *J. Econ. Theory* 17 (February 1978): 79–98.

Dulles, Eleanor L. *The French Franc, 1914–1928: The Facts and Their Interpretation.* New York: Macmillan, 1929.

Einzig, Paul. *The Theory of Forward Exchange.* London: Macmillan, 1937.

Flood, Robert P., and Garber, Peter M. "An Economic Theory of Monetary Reform." *J.P.E.* 88, no. 1 (February 1980): 24–58.

Frenkel, Jacob A. "The Forward Exchange Rate, Expectations, and the Demand for Money: The German Hyperinflation." *A.E.R.* 67 (September 1977): 653–70.

———. "Purchasing Power Parity: Doctrinal Perspective and Evidence from the 1920s." *J. Internat. Econ.* 8 (May 1978): 169–91.

———. "Further Evidence on Expectations and the Demand for Money during the German Hyperinflation." *J. Monetary Econ.* 5 (January 1979): 81–97.

Frenkel, Jacob A., and Clements, Kenneth W. "Exchange Rates in the 1920's: A Monetary Approach." Working Paper no. 290, Nat. Bur. Econ. Res., 1978.

Fuller, Wayne A. *Introduction to Statistical Time Series.* New York: Wiley, 1976.

Geweke, John. "The Relative Asymptotic Strength of Some Tests Used in Time Series Analysis." Working Paper, Univ. Wisconsin—Madison, Soc. Systems Res. Inst., 1979.

Geweke, John, and Feige, Edgar. "Some Joint Tests of the Efficiency of Markets for Forward Foreign Exchange." *Rev. Econ. and Statis.* 61 (August 1979): 334–41.

Grauer, Frederick L. A.; Litzenberger, Robert H.; and Stehle, Richard E. "Sharing Rules and Equilibrium in an International Capital Market under Uncertainty." *J. Financial Econ.* 3 (June 1976): 233–56.

Hansen, Lars Peter. "The Asymptotic Distribution of Least Squares Estimators with Endogenous Regressors and Dependent Residuals." Working Paper, Carnegie-Mellon Univ., Graduate School Indus. Admin., 1979.

Kareken, John, and Wallace, Neil. "Samuelson's Consumption-Loan with Country-specific Fiat Monies." Res. Dept. Staff Report 24, Federal Reserve Bank of Minneapolis, 1978.

Leamer, Edward E. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: Wiley, 1978.
Levich, Richard M. "Further Results on the Efficiency of Markets for Foreign Exchange." Paper presented at a conference of the Federal Reserve Bank of Boston, Bald Peak, N.H., October 1978.
Lucas, Robert E., Jr. "Asset Prices in an Exchange Economy." *Econometrica* 46 (November 1978): 1429–45.
McCulloch, J. Huston. "Operational Aspects of the Siegel Paradox: Comment." *Q.J.E.* 89 (February 1975): 170–74.
Nickelsburg, G. "Flexible Exchange Rates and Uncertain Government Policy: A Theoretical and Empirical Analysis." Ph.D. dissertation, Univ. Minnesota, 1979.
Roll, Richard, and Solnik, Bruno. "A Pure Foreign Exchange Asset Pricing Model." *J. Internat. Econ.* 7 (May 1977): 161–79.
Roper, Don E. "The Role of Expected Value Analysis for Speculative Decisions in the Forward Currency Market: Comment." *Q.J.E.* 89 (February 1975): 157–69.
Rozanov, Yu A. *Stationary Random Processes*. Translated by A. Feinstein. San Francisco: Holden-Day, 1967.
Sargent, Thomas J. *Macroeconomic Theory*. New York: Academic Press, 1979.
Siegel, Jeremy J. "Risk, Interest, and Forward Exchange." *Q.J.E.* 86 (May 1972): 303–9.
Sims, Christopher. "Least Squares Estimation of Autoregressions with Some Unit Roots." Working Paper, Univ. Minnesota, Dept. Econ., 1978.
Stockman, Alan C. "Risk, Information, and Forward Exchange Rates." In *The Economics of Exchange Rates: Selected Studies*, edited by Jacob A. Frenkel and Harry G. Johnson. Reading, Mass.: Addison-Wesley, 1978.