# Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability

By NELSON C. MARK*

*Regressions of multiple-period changes in the log exchange rate on the deviation of the log exchange rate from its "fundamental value," display evidence that long-horizon changes in log nominal exchange rates contain an economically significant predictable component. To account for small-sample bias and size distortion in asymptotic tests, inference is drawn from bootstrap distributions generated under the null hypothesis that the log exchange rate is unpredictable. The bias-adjusted slope coefficients and $R^2$'s increase with the forecast horizon, and the out-of-sample point predictions generally outperform the driftless random walk at the longer horizons. (JEL F31)*

The difficulty in predicting the logarithm of exchange rates has been a longstanding problem in international economics. In a seminal paper, Richard Meese and Kenneth Rogoff (1983a) conduct a monthly postsample fit analysis from 1976:11 to 1981:6 for U.S. dollar prices of the deutsche mark, the pound, and the yen.[1] Point predictions of the driftless random walk dominate those of their regressions for all three currencies at

*Department of Economics, Ohio State University, 410 ARPS Hall, Columbus, OH 43210. This is a revised version of a paper previously circulated under the title, "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability and Overshooting." I thank Joshua Aizenman, Stephen Cecchetti, Stephen Cosslett, Robert Cumby, Francis Diebold, Charles Engel, John Huizinga, G. S. Maddala, Frederic Mishkin, Pierre Perron, and seminar participants at the · International Economics Workshop Conference at the University of Chicago, the University of Pennsylvania, Penn State University, the Université de Montréal, Rutgers University, Vanderbilt University, the University of Virginia, the University of Washington, and the National Bureau of Economic Research for useful comments on earlier drafts. The comments of two anonymous referees helped improve the paper. I bear sole responsibility for any errors.

[1] They examine flexible-price monetary models with and without current-account effects, and a sticky-price monetary model. Their out-of-sample predictions are generated from regressions with coefficients estimated with in-sample data but postsample realizations of the regressors, hence the term postsample "fit."

the 6- and 12-month horizons, and for two of the three currencies at the 1-month horizon. In Meese and Rogoff (1988), changes in log real exchange rates are regressed on real interest-rate differentials to forecast log real dollar prices of the deutsche mark, the pound, the yen, and the implied cross rates, at 1-, 6-, and 12-month horizons. Over the 1980:11–1986:3 period (1980:11–1985:1 for yen rates) they find that forecasts from the random walk have lower root-mean-square error (RMSE) than those from their regressions in 32 of 36 postsample fit experiments. Furthermore, on the basis of asymptotic tests, the regression forecasts are never significantly better than the driftless random walk.

Work that attempts to exploit nonlinearities in the exchange-rate process to predict future log exchange rates has similarly been limited in its success. Charles Engel and James Hamilton (1990) analyze quarterly out-of-sample point predictions of their segmented-trends model from 1984 to 1988 for the deutsche mark, the French franc, and the pound. They find that forecasts generated by their model at the 4-quarter horizon for the deutsche mark and French franc rates are outperformed by the random walk with drift. Similarly, Francis Diebold and James Nason (1990) examine ten exchange rates and find that forecasts from nonparametric estimates of the conditional

mean weekly percentage change in the nominal exchange rate generally cannot improve upon the random walk without drift during the 66 weeks beginning in 1986 and ending with week 38 of 1987.[2]

This paper presents evidence that long-horizon changes in the logarithm of spot exchange rates are predictable. The currency values that I study are end-of-quarter U.S. dollar prices of the Canadian dollar, the deutsche mark, the Swiss franc, and the yen from 1973 to 1991, and the evidence is based on estimated projections of 1-, 4-, 8-, 12-, and 16-quarter changes in the log exchange rate on the deviation of the current log exchange rate from its "fundamental value." This fundamental value is motivated by a monetary model of exchange-rate determination and is defined to be a linear combination of log relative money stocks and log relative real incomes. Examination of the estimated slope coefficients, regression $R^2$'s, and out-of-sample point predictions relative to those of the driftless random walk indicate that an economically significant predictable component is contained in the Swiss franc and yen rates at each of the horizons considered, and for the deutsche mark at 12 and 16 quarters.

Statistical inference is carried out by testing whether the regression slope coefficient is zero, and by testing whether the out-of-sample point predictions yield significant improvements over those of the driftless random walk. This task, however, is made complicated by the unavailability of a sufficiently long time series over the float. One complication is that under the null hypothe-

sis that the log exchange rate is unpredictable, the least-squares estimator is biased away from zero in small samples. This problem arises because the independent variable is predetermined but not exogenous. A second complication is that the asymptotic tests have the wrong size even when small-sample bias has been taken into account. This problem arises in part because the observations are sampled quarterly but the $k$-quarter change in the log exchange rate is used as the dependent variable which induces $(k-1)$th-order serial correlation into the regression error under the null. When the overlap is large relative to the sample size, the asymptotic distribution provides a poor approximation to the exact distribution.

To account for bias and size distortion, I draw inference from Gaussian and nonparametric bootstrap distributions generated under the null hypothesis that the log exchange rate is unpredictable. Under the nonparametric bootstrap, the hypothesis that the slope coefficients across the five horizons are jointly zero can be rejected at the 5-percent level for the deutsche mark and the Swiss franc, while the marginal significance level of this joint test statistic is 0.41 for the Canadian dollar and 0.26 for the yen. The hypothesis that the regression and the random walk generate equally accurate out-of-sample forecasts at all five forecast horizons can also be rejected at the 5-percent level for the deutsche mark and Swiss franc. (Similar inference is drawn from the parametric bootstrap.) These statistical results are somewhat mixed, but it is worth keeping in mind that the inability of the tests to provide a blanket rejection of the null hypothesis at the conventional 5-percent level may well be due to the shortness of the available time series. At the 16-quarter horizon, the data set contains fewer than five nonoverlapping observations while the out-of-sample forecast experiment generates fewer than two nonoverlapping observations. Consequently, tests with 5-percent size may be excessively cautious in guarding against type-I errors and might reasonably be adjusted upward. To facilitate such an adjustment, marginal significance

---

[2] Wing Woo (1985) reports that the monthly out-of-sample fit of his monetary model attains an RMSE reduction in excess of 50 percent relative to the random walk over the 1980:3–1981:10 period. His results appear to be driven in part by the fact that the data he employs are deviations from a quadratic trend that was estimated over the entire sample. This procedure has two implications. First, postsample information is exploited in the prediction exercise. Second, and more importantly, the random walk is obviously a bad candidate for forecasting residuals from a quadratic trend, as these data have had their low-frequency components filtered out.

levels of tests are reported throughout the paper.

My out-of-sample forecast analysis begins in 1981 so as to include the 1985 and 1987 turning points of the dollar. At the 1-quarter horizon, the regression point predictions for the Canadian dollar, the Swiss franc, and the yen have RMSE's that are slightly lower than those of the random walk. At the 4-quarter horizon the regressions generate a reduction in prediction RMSE, relative to the random walk, of 2 percent for the Swiss franc and 7 percent for the yen. Although direct comparisons to the literature quickly become tangled due to discrepancies in the sample periods and the exchange rates studied by various authors, my out-of-sample point predictions at the 1- and 4-quarter horizons are generally consistent with findings from earlier research that the log exchange rate is difficult to predict. The situation changes markedly, however, when long horizons are examined.[3] At 16 quarters, the RMSE's of my regression forecasts relative to the random walk are 0.52, 0.41, and 0.58 for the deutsche mark, the Swiss franc, and the yen, respectively.[4]

The regressions also display a positive relationship among the slope coefficients, the $R^2$'s, and the forecast horizon. For example, using the log U.S. dollar/deutsche-mark rate, the bias-adjusted slope coefficient increases from 0.01 to 1.02, and the bias-adjusted $R^2$ increases from 0.01 to 0.64 when the horizon is lengthened from 1 to 16 quarters. A similar relationship among the slope coefficients, $R^2$'s, and horizon, has been found by John Campbell and Robert Shiller (1988) and Eugene Fama and Kenneth French (1988), who regress stock returns on dividend yields. Their results have been interpreted by some authors as evidence that irrational agents cause stock prices to overreact to changes in dividends. While the results presented below are consistent either with overreaction or slow adjustment of the exchange rate to changes in economic fundamentals, these gradual adjustments need not imply irrationality. Instead, they may be the consequence of nominal rigidities as emphasized by Rudiger Dornbusch (1976), Jeffrey Frankel (1979), and Michael Mussa (1982).

The remainder of the paper proceeds as follows. The next section motivates the regressions that I estimate. The data are described in Section II. Econometric issues, the out-of-sample prediction exercise, and the bootstrap procedures used to conduct inference are discussed in Section III. This section also describes a limited sensitivity analysis that I perform.[5] Section IV contains the empirical results, and concluding remarks are contained in Section V.

## I. Motivating the Long-Horizon Regressions

Let $e_t$ be the date-$t$ logarithm of the domestic-currency price of one unit of foreign exchange, let $m_t$ be the date-$t$ logarithm of the domestic money stock, let $y_t$ be the date-$t$ logarithm of domestic real in-

---

[3]Recent empirical research has uncovered many features of exchange-rate behavior that suggest the presence of a predictable component. John Huizinga (1987) reports that real exchange rates from the modern float display mean reversion, and Robert Cumby and Huizinga (1991) find that real exchange rates are correlated with past exchange-rate changes and past inflation differentials, while Marianne Baxter (1993) finds that log real exchange rates are correlated with real interest differentials at business-cycle frequencies. In addition, Mark Rush and Steven Husted (1985), Hali Edison (1987), Niso Abauf and Philippe Jorion (1990), Diebold et al. (1991), Eric Fisher and Joon Park (1991), Vittorio Grilli and Graciela Kaminsky (1991), and Yoonbai Kim (1990) find evidence that long-run purchasing-power parity holds.

[4]Previous research has largely confined its attention to horizons of one year or less. Meese and Rogoff (1983b) provide the exception, but they report out-of-sample fit results in which the parameters are not estimated but are fixed a priori.

[5]The main conclusions of the paper are not sensitive to various modifications to the data generating process employed in the bootstrapping procedure. The results also do not appear to hinge on my use of quarterly data. A sensitivity analysis and estimation results for the Canadian dollar, the deutsche mark, and the yen using monthly data (monthly data for Switzerland were not available) are contained in an appendix that is available from the author upon request.

come, and let stars (*) denote quantities of the foreign country. Now consider the simplest monetary model of exchange-rate determination originally popularized by Jacob Frenkel (1976), Mussa (1976), and John Bilson (1978), and used recently in target-zone analyses developed by Paul Krugman (1991). In this model, it is assumed that purchasing-power parity (PPP) and uncovered-interest parity (UIP) hold, and that log money demand depends linearly and contemporaneously on log real income and the nominal interest rate. Denote the common (to the home and foreign country) money-demand income elasticity by $\lambda$ $(0 \leq \lambda \leq 1)$ and the common money-demand interest semi-elasticity by $\phi \geq 0$. The log exchange-rate solution is the present-value relation

$$(1) \qquad e_t = \frac{\delta}{\phi} E_t \left( \sum_{j=0}^{\infty} \delta^j f_{t+j} \right) + c$$

where

$$(2) \qquad f_t \equiv (m_t - m_t^*) - \lambda(y_t - y_t^*)$$

is the date-$t$ fundamental for this model, $\delta = \phi/(1+\phi)$ and $c$ is a constant arising from the base-year problem in the PPP relationship.

As a benchmark, let $\{f_t\}$ follow a driftless random walk. Then (1) reduces to the exact relation

$$(3) \qquad e_t = f_t + c.$$

The benchmark model implies a direct relationship between $\{e_t\}$ and $\{f_t\}$, and it inspires references to $f_t$ as the log exchange rate's fundamental value.

It is well known that deviations of the log exchange rate from the fundamentals are common and persistent since both casual observation and formal econometric studies reject this simple monetary model.[6] However, a reasonable hypothesis to put forth is that the log exchange rate returns to its fundamental value over time so that its current state of evolution can be characterized by deviations from the benchmark monetary model. To investigate this idea, project the $k$-period-ahead change in the log exchange rate on its current deviation from the fundamental value to obtain

$$(4) \qquad e_{t+k} - e_t = \alpha_k + \beta_k z_t + \nu_{t+k,t}$$

where $z_t \equiv f_t - e_t$, $\alpha_k$, and $\beta_k$ are the linear least-squares projection coefficients, and $\nu_{t+k,t}$ is the projection error. When $e_t$ is below its fundamental value, it is expected to rise over time, which implies that the slope coefficient, $\beta_k$, should be positive and should initially increase with the horizon $k$. Throughout the paper, I employ a fundamental value that is constructed with a fixed value of $\lambda = 1$.[7]

It is worth emphasizing that the goal of this paper is to conduct an empirical examination of (4) to determine the predictive content of the state variable $z_t$. I do not estimate a well-articulated model of exchange-rate dynamics, and I remain agnostic about using (4) to test the implications of a particular theory.

## II. The Data

The data are quarterly observations for the United States, Canada, Germany, Japan, and Switzerland and were obtained from the OECD *Main Economic Indicators*. The sample consists of 76 observations extend-

---

[6] Formal econometric studies systematically reject the naive monetary model. For example, Woo (1985) and Robert Driskill et al. (1992) find strong evidence of lagged adjustment in the exchange-rate solution. Kenneth West (1987) argues that either large deviations from PPP or large money-demand errors are required to save the monetary model, Richard Meese (1986) rejects the monetary model in favor of a bubbles speci-

fication, and Meese and Kenneth Singleton (1983) document violations of variance-bounds restrictions implied by the model. At a very basic level, both PPP and UIP are well known to be violated in the data.

[7] To the extent that the results are robust to variations in $\lambda$, monetary factors are relatively more important than relative real income in predicting exchange-rate movements. In fact, the patterns that the estimates display with respect to $k$ are robust to variations of $\lambda$ from 0 to 1, but the statistical significance of the results is weaker for values near zero. I report results using $\lambda = 0.0, 0.1, 0.5,$ and $0.9$ in the unpublished appendix (available from the author upon request).

ing from 1973:2 to 1991:4. The sample includes Germany and Japan because they, along with the United States, are large and important economies. I include the Swiss franc because Switzerland has never participated in monetary arrangements such as the Snake (the arrangement of exchange rates among European Economic Community countries during 1972–1979) or the European Monetary System and has allowed the franc to float freely. Finally, I include the Canadian dollar because of Canada's close economic ties with the United States.

The exchange rates are U.S. dollar prices of the foreign currency. The monetary variable used to construct the fundamental value in the deutsche mark, the Swiss franc, and the yen regressions is M1 (currency plus checkable deposits). Canadian M1 was not used because the log exchange-rate's deviation from the M1-constructed fundamental displays a pronounced upward trend throughout the sample and appears to be nonstationary. As a result, M3 (M1 plus quasi-money) serves as Canadian money in these regressions.

Real income is measured by quarterly real GNP for the United States and quarterly real GDP for the remaining countries. These real income data are seasonally adjusted, while the money data are not. To remove the seasonality in money, the monetary variable is formed by summing the current observation on money with three lagged values.

### III. Econometric Methodology

Subsection III-A describes asymptotic covariance matrix estimators that are robust to the serial correlation induced into the regression error when the forecast horizon extends beyond the sampling interval. I employ these estimators to construct $t$ ratios for testing whether the regression slope coefficient is zero and to construct a statistic proposed by Diebold and Robert Mariano (1993) to test whether the regression and the random walk generate equally accurate forecasts. The out-of-sample forecast experiment and Diebold and Mariano's statistic are described in Subsection III-B. As mentioned in the Introduction, because the sample is relatively small and because the extent of the observation overlap is large relative to the sample size, ordinary least-squares (OLS) analysis is biased under the null hypothesis, and asymptotic tests suffer from size distortion. Subsection III-C describes the nature of the OLS bias while Subsection III-D describes the construction of the Gaussian and the nonparametric bootstrap distributions that are used to estimate the small-sample bias and to conduct inference.

### A. Asymptotic Covariance Matrices

Extending the forecast horizon beyond the sampling interval $(k > 1)$, induces $(k-1)$th-order serial correlation in the error, $\nu_{t+k,t}$, of the regression (4) under the null hypothesis. Consistent estimation of the least-squares asymptotic covariance matrix thus requires estimating the spectral density matrix of $\boldsymbol{\eta}_{t+k} = (\nu_{t+k,t}, \nu_{t+k,t} z_t)'$ at frequency zero. For this purpose, I employ two versions of Whitney Newey and Kenneth West's (1987) covariance matrix estimator.[8] The versions differ according to the rule used to determine the truncation lag for the Bartlett window.

The first version arbitrarily sets the truncation lag at 20. Because the longest forecast horizon that I entertain is 16 periods, 20 lags should do a reasonable job of accounting for the serial correlation of the error term. The second version employs a data-dependent formula provided by Donald Andrews (1991) that takes the esti-

---

[8]Constructing the bootstrap distribution with the covariance matrix estimator proposed by Lars Hansen and Hodrick (1980), which uses uniform weights and $k-1$ autocovariances, proved to be problematic and was not used because of a preponderance of trials in which this estimator was not positive definite. The estimator that I do use, however, is not exactly that of Newey and West (1987) since in estimating the $j$th-order autocovariance matrix of $\{\eta_t\}$, I make a degrees-of-freedom correction by dividing $\Sigma \boldsymbol{\eta}_t \boldsymbol{\eta}'_{t+j}$ by $T-j$ instead of $T$. It was pointed out to me by Kenneth West (pers. comm.) that the Newey and West (1987) estimator does not make a degrees-of-freedom correction and that positive definiteness is not guaranteed when such a correction is made.

mates from a parametric model of $\{\eta_{t+k}\}$ as inputs. The elements of $\{\eta_{t+k}\}$ follow an MA($k-1$) process under the null hypothesis, but I do not use the MA model as it is not amenable to the bootstrap analysis. Instead, to maintain computational feasibility, I employ a univariate AR(1) as an approximating model.

The rule is implemented as follows. Let $\{\hat{\rho}_j, \hat{\sigma}_j^2 : j = 1,2\}$ denote the estimated autocorrelation coefficient and innovation variance from the AR(1) processes for the elements of $\{\eta_{t+k}\}$. Then the truncation lag, $A$, for the Bartlett window in Newey and West's estimator given by Andrews's rule is $A = 1.1447[\hat{a}(1)T]^{1/3}$, where

$$\hat{a}(1) = \frac{\displaystyle\sum_{j=1}^{2} \frac{4\hat{\rho}_j^2 \hat{\sigma}_j^4}{\left(1 - \hat{\rho}_j\right)^6 \left(1 + \hat{\rho}_j\right)^2}}{\displaystyle\sum_{j=1}^{2} \frac{\hat{\sigma}_j^4}{\left(1 - \hat{\rho}_j\right)^4}}.$$

### B. Out-of-Sample Prediction

The out-of-sample forecasts for a given horizon $k$ are constructed by running the regression (4) with data up through date $t_0 < T$, so that the last observation used is $(e_{t_0} - e_{t_0-k}, z_{t_0-k})$. Let $\hat{\alpha}(k; t_0)$ and $\hat{\beta}(k; t_0)$ be the coefficients estimated with these data. The first $k$-horizon forecast is

$$(5) \quad \hat{e}_{t_0+k} - e_{t_0} = \hat{\alpha}(k; t_0) + \hat{\beta}(k; t_0) z_{t_0}.$$

Next, the time subscript is advanced, and the procedure is repeated for $t_0 + 1$, $t_0 + 2, \ldots, T - k$. Forecast accuracy is determined by the square-error criteria, and one measure that I employ is the ratio of the regression's prediction RMSE relative to that of the random walk. This statistic is denoted by (OUT/RW).

To employ Diebold and Mariano's (1993) test of the hypothesis that the forecasts from two competing models are equally accurate, denote the date-$t$ forecast error of model $i$ ($i = 1,2$) by $u_{i,t}$, the number of forecasts by $N_f = T - t_0 - k + 1$, the sample mean-square

error differential by

$$\bar{d} = N_f^{-1} \Sigma_{t=t_0+k}^{T} \left(u_{1,t}^2 - u_{2,t}^2\right)$$

and the spectral density of $(u_{1,t}^2 - u_{2,t}^2)$ at frequency 0 by $f_d(0)$. The test is based on the statistic

$$(6) \quad \mathcal{DM} = \frac{\bar{d}}{\sqrt{\dfrac{2\pi \hat{f}_d(0)}{N_f}}}.$$

where $\hat{f}_d(0)$ is a consistent estimate of $f_d(0)$. Under the null hypothesis of equal forecast accuracy, the mean-square error differential is zero, and $\mathcal{DM}$ has an asymptotic standard normal distribution. A consistent estimate of the spectral density at frequency zero is obtained using the method of Newey and West (1987). Both a fixed lag of 20 and Andrew's AR(1) approximating rule for setting the truncation lag are used. Andrews's rule is implemented by fitting an AR(1) to the scalar quantity $(u_{1,t}^2 - u_{2,t}^2)$, where $\hat{a}(1) = 4[\hat{\rho}/(1-\hat{\rho})(1+\hat{\rho})]^2$, and $\hat{\rho}$ is the estimated first-order autocorrelation coefficient.

### C. Small-Sample Bias

Robert Stambaugh (1986) shows that the least-squares estimator of $\beta_k$ may be biased in small samples. He discusses a regression analogous to (4) with $k = 1$, in which the regressor, $\{z_t\}$, follows an AR(1).[9] In the present context, that system is

$$(7) \quad e_{t+1} - e_t = \alpha_1 + \beta_1 z_t + \varepsilon_{1,t+1}$$

$$(8) \quad z_{t+1} = \delta + \gamma z_t + \varepsilon_{2,t+1}$$

where $\{(\varepsilon_{1,t}, \varepsilon_{2,t})'\}$ is an independently and identically distributed vector sequence. As is well known, if $\{z_t\}$ follows an AR(1), the least-squares estimator of $\gamma$ in a sample of $T$ observations is biased toward zero with

---

[9] Robert Hodrick (1992) and Charles Nelson and Myung Kim (1992) study bias and size-distortion issues in regressions of equity returns on dividend yields.

the bias given by

$$(9) \qquad E(\hat{\gamma} - \gamma) \simeq - \left( \frac{1 + 3\gamma}{T} \right).$$

Stambaugh shows that the bias in the least-squares estimator of $\beta_1$ is proportional to the bias in $\hat{\gamma}$ where the factor of proportionality is the slope coefficient in a regression of $\varepsilon_{1,t}$ on $\varepsilon_{2,t}$,

$$(10) \quad E\left(\hat{\beta}_1 - \beta_1\right)$$

$$= \left( \frac{\mathrm{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t})}{\mathrm{Var}(\varepsilon_{2,t})} \right) E(\hat{\gamma} - \gamma).$$

From (9) it can be seen that the bias in $\hat{\gamma}$ can be large for small $T$. Although the bias described above pertains to $k = 1$ and an AR(1) regressor, I will show that the results from bootstrapping the slope coefficient suggest that this same intuition carries over to $k > 1$ and for regressors that follow more complex time-series processes.

### D. Bootstrap Distributions

The bootstrap distributions are built upon estimated values of the restricted vector autoregression (VAR) that embodies the null hypothesis that the exchange rate is unpredictable:[10]

$$(11) \qquad \Delta e_t = a_0 + \varepsilon_{1,t}$$

$$(12) \qquad z_t = b_0 + \sum_{j=1}^{p} b_j z_{t-j} + \varepsilon_{2,t}.$$

Let $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$, let $V = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$, and let the estimates be $(\hat{a}_0, \hat{b}_0, \hat{b}_1, \dots, \hat{b}_4, \hat{V})$. The Gaussian bootstrap distribution is generated by running 2,000 trials where each trial $(i = 1, \dots, 2{,}000)$ proceeds as follows:

1. Draw a vector sequence of observations $\{\underline{\varepsilon}_t^i\}_{t=1}^{n+T}$ from a bivariate normal distribution with mean 0, covariance matrix $\hat{V}$, $n = 1{,}000$, and $T = 76$.

[10] The design of the null hypothesis is similar to that used by Hodrick (1992) in his study of stock-return prediction regressions.

2. Generate sequences of observations $\{\Delta e_t^i\}_{t=1}^{n+T}$, $\{z_t^i\}_{t=1}^{n+T}$ of length $n + T$ according to

$$\Delta e_t^i = \hat{a}_0 + \varepsilon_{1,t}^i$$

$$z_t^i = \hat{b}_0 + \sum_{j=1}^{p} \hat{b}_j z_{t-j}^i + \varepsilon_{2,t}^i$$

using the sample mean for the initial values, $(z_0, z_{-1}, z_{-2}, z_{-3})$, in starting up the autoregression.

3. Drop the first $n$ observations and use the remaining $T$ data points to regress $e_{t+k}^i - e_t^i$ on $z_t^i$ to obtain $\hat{\beta}_k^i$, $t_k^i(20)$ and $t_k^i(A)$ (the asymptotic $t$ ratios with truncation lags set at 20 and by the Andrews rule, respectively), and $R_{k,i}^2$ (the $R^2$ from this regression), for each horizon, $k$. Also, perform the out-of-sample prediction exercise on these simulated observations and calculate $(\mathrm{OUT/RW})_k^i$, $\mathcal{DM}_k^i(20)$, and $\mathcal{DM}_k^i(A)$ (the regression forecast RMSE relative to that of the random walk, and the Diebold-Mariano statistics with truncation lags set at 20 and set by the Andrews rule).

These 2,000 observations of $\hat{\beta}_k^i$, $t_i(20)$, $t_i(A)$, $R_{k,i}^2$, $(\mathrm{OUT/RW})_k^i$, $\mathcal{DM}_k^i(20)$, and $\mathcal{DM}_k^i(A)$ form the Gaussian bootstrap distribution for the slope coefficient, the asymptotic $t$ ratios, the $R^2$, the $(\mathrm{OUT/RW})$ statistics, and the $\mathcal{DM}$ statistics under the null hypothesis that the log exchange rate is unpredictable.

To investigate whether the size-adjustment procedure is sensitive to the normality assumption, I also generate nonparametric bootstrap distributions. These distributions are built up as described above, but with the following modifications to steps 1 and 2. In step 1, I set $n = 0$ and generate innovation sequences of length $T$ by random resampling with replacement of the fitted residuals. In step 2, the initial values used to start up the autoregression are randomly drawn using the four previous observations on $z_t$ corresponding to the first draw of the residual, $\varepsilon_{2,1}^i$.

## IV. Empirical Results

Subsection IV-A reports estimates of the model under the null hypothesis that serves as the data generating process (DGP) in building the bootstrap distribution. A matter of potential concern is that the DGP is estimated with error. A sensitivity analysis conducted to address this issue is also described in this section. Subsection IV-B reports regression results for forecast horizons at $k = 1, 4, 8, 12$, and 16 quarters using data over the full sample. The results of the out-of-sample prediction experiment are reported in Subsection IV-C.

### A. Estimates and Properties of the Estimated DGP

Table 1 displays estimation results for the system in equations (11) and (12). I also implement Lagrange multiplier (LM) tests for first- and fourth-order residual serial correlation and for fourth-order autoregressive conditional heteroscedasticity (ARCH). The LM statistic for serial correlation is $TR^2$ from a regression of the residual on the regressors and either one or four lagged residuals. The LM test of Robert Engle (1982) is employed as the test for ARCH residuals. From panel A of the table, there is little evidence of serial correlation or ARCH in the log exchange-rate differences, as none of the chi-square statistics is significant at the 10-percent level.[11]

Panel B displays the estimates of the process governing the regressor $\{z_t\}$. For the deutsche mark, an AR(5) is required to achieve serially uncorrelated residuals while AR(4)'s are adequate for the other three currencies. The LM tests produce little evidence for ARCH. Panel B also reports a number of additional features of the estimated process. The next line in the table

displays augmented Dickey-Fuller (ADF) statistics calculated with a constant and $p - 1$ lags. Since the 10-percent critical value is $-2.58$, these tests cannot reject the hypothesis that $\{z_t\}$ contains a unit root. Their inability to reject, however, is not sufficient cause to abandon the stationarity assumption since, as Stephen Blough (1992) has shown, the power of generic unit-root tests is bounded above by the size of the test. Nevertheless, the $\{z_t\}$ processes are quite persistent. The largest root implied by the estimated autoregressive polynomial ranges in modulus from 0.78 for the Swiss franc to 0.92 for the Canadian dollar while the lower $(\rho_1)$ and upper $(\rho_u)$ 95-percent confidence bands are large and contain 1 for each currency. These bands and the median estimate $(\rho_m)$ of the largest root are obtained using table A.1 of James Stock (1991), which tabulates the correspondence between the ADF statistic and the confidence band of the largest root.

A potential concern is that estimates of the largest root are downward-biased even when its true value is less than 1. Indeed, by comparing the largest root implied by the estimated autoregressive polynomial to the median estimate, it is seen that downward bias appears to be important for the deutsche mark, the Swiss franc, and the yen. To investigate whether inferences drawn from the bootstrap distributions are particularly sensitive to misspecification of the largest root, I produced alternative size adjustments as follows. First, I altered the magnitude of the largest root of the autoregressive polynomial in the data generating process (DGP). Then, holding the values of the remaining roots constant at their estimated values, I determined the implied AR coefficients and used them as the DGP parameters. In one set of experiments, I set the largest root equal to the median value, $\rho_m$. In a second set of experiments, the largest root is simply increased by 5 percent, which approximately corrects for the bias implied by (9) for a root near 1 and $T = 76$. The results obtained by these modifications to the DGP do not change the main conclusions of the paper and are contained in the unpublished appendix (available from the author upon request).

---

[11]The absence of ARCH at quarterly horizons is not surprising. Richard Baillie and Tim Bollerslev (1989) find evidence of ARCH in exchange-rate changes at short (daily) intervals, but the strength of this evidence decreases as their sampling interval lengthens. At monthly horizons, they find virtually no evidence of ARCH.

TABLE 1—ESTIMATES OF THE DATA GENERATING PROCESS UNDER
THE NULL HYPOTHESIS

A.

$$\Delta e_t = a_0 + \varepsilon_{1,t}$$

| Statistic | Canadian dollar | Deutsche mark | Swiss franc | Yen |
|---|---|---|---|---|
| $a_0$ | −0.002 | 0.006 | 0.011 | 0.010 |
| (SE) | (0.003) | (0.007) | (0.008) | (0.007) |
| LM tests for first- and fourth-order residual serial correlation: | | | | |
| $X^2_{[1]}$ | 0.462 | 1.033 | 0.604 | 1.988 |
| (MSL) | (0.497) | (0.309) | (0.437) | (0.159) |
| $X^2_{[4]}$ | 3.949 | 7.572 | 3.518 | 3.138 |
| (MSL) | (0.413) | (0.109) | (0.475) | (0.535) |

B.

$$z_t = b_0 + \sum_{j=1}^{p} b_j z_{t-j} + \varepsilon_{2,t}$$

| Statistic | Canadian dollar | Deutsche mark | Swiss franc | Yen |
|---|---|---|---|---|
| $b_0$ | 0.067 | −0.255 | 0.197 | −0.008 |
| (SE) | (0.030) | (0.149) | (0.080) | (0.007) |
| $b_1$ | 1.224 | 1.085 | 1.113 | 1.105 |
| (SE) | (0.110) | (0.117) | (0.118) | (0.120) |
| $b_2$ | −0.260 | −0.217 | −0.247 | −0.167 |
| (SE) | (0.179) | (0.172) | (0.177) | (0.179) |
| $b_3$ | 0.405 | 0.260 | 0.219 | 0.054 |
| (SE) | (0.178) | (0.173) | (0.179) | (0.179) |
| $b_4$ | −0.417 | 0.118 | −0.190 | −0.072 |
| (SE) | (0.106) | (0.178) | (0.117) | (0.119) |
| $b_5$ | — | −0.307 | — | — |
| (SE) | — | (0.124) | — | — |
| LM tests for first- and fourth-order residual serial correlation: | | | | |
| $X^2_{[1]}$ | 0.404 | 0.576 | 0.890 | 0.227 |
| (MSL) | (0.525) | (0.448) | (0.346) | (0.634) |
| $X^2_{[4]}$ | 5.505 | 2.757 | 4.732 | 2.997 |
| (MSL) | (0.239) | (0.599) | (0.316) | (0.558) |
| LM test for fourth-order residual ARCH: | | | | |
| $X^2_{[4]}$ | 4.638 | 2.462 | 1.512 | 3.890 |
| (MSL) | (0.326) | (0.651) | (0.825) | (0.421) |
| Augmented Dickey-Fuller statistics for $\{z_t\}$, $p-1$ lags: | | | | |
| ADF | −2.176 | −1.597 | −2.431 | −1.851 |
| Implied largest root | 0.917 | 0.905 | 0.780 | 0.879 |
| Largest-root 95-percent confidence interval $(\rho_l, \rho_u)$ and median $(\rho_m)$: | | | | |
| $\rho_l$ | 0.777 | 0.864 | 0.745 | 0.822 |
| $\rho_m$ | 0.917 | 0.997 | 0.891 | 0.956 |
| $\rho_u$ | 1.038 | 1.052 | 1.034 | 1.046 |

*Notes:* Chi-square statistics are for the $TR^2$ Lagrange multiplier test for first- and fourth-order residual serial correlation and fourth-order ARCH. MSL refers to the marginal significance levels for a one-tail test. The 10-percent critical value for the ADF test is −2.58. For the last rows in the table, $\rho_l$ and $\rho_u$ are the 95-percent confidence intervals for the largest root implied by the ADF statistic from Stock's (1991) distribution, and $\rho_m$ is the median.

TABLE 2—REGRESSION ESTIMATES AND BOOTSTRAP DISTRIBUTIONS

| (i) | (ii) | (iii) | (iv) | (v) | (vi) | (vii) | (viii) | (ix) | (x) | (xi) | (xii) | (xiii) | (xiv) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $k$ | $\hat{\beta}_k$ | Adj-p | Adj-n | $R^2$ | Adj-p | Adj-n | $t_k(20)$ | MSL-p | MSL-n | $t_k(A)$ | $A$ | MSL-p | MSL-n |
| Canadian dollar: | | | | | | | | | | | | | |
| 1 | 0.040 | 0.027 | 0.028 | 0.059 | 0.050 | 0.051 | 3.051 | 0.076 | 0.077 | 2.172 | 1 | 0.065 | 0.066 |
| 4 | 0.155 | 0.109 | 0.114 | 0.179 | 0.144 | 0.146 | 2.389 | 0.186 | 0.195 | 2.168 | 12 | 0.172 | 0.175 |
| 8 | 0.349 | 0.264 | 0.264 | 0.351 | 0.287 | 0.285 | 2.539 | 0.222 | 0.230 | 2.527 | 19 | 0.201 | 0.216 |
| 12 | 0.438 | 0.320 | 0.315 | 0.336 | 0.251 | 0.235 | 1.961 | 0.317 | 0.343 | 1.936 | 29 | 0.323 | 0.343 |
| 16 | 0.450 | 0.295 | 0.287 | 0.254 | 0.146 | 0.121 | 1.542 | 0.420 | 0.447 | 1.512 | 33 | 0.436 | 0.456 |
| Deutsche mark: | | | | | | | | | | | | | |
| 1 | 0.035 | 0.012 | 0.016 | 0.015 | 0.005 | 0.006 | 1.836 | 0.280 | 0.252 | 0.929 | 2 | 0.408 | 0.391 |
| 4 | 0.205 | 0.114 | 0.126 | 0.104 | 0.065 | 0.068 | 2.902 | 0.169 | 0.157 | 2.290 | 15 | 0.206 | 0.193 |
| 8 | 0.554 | 0.380 | 0.410 | 0.265 | 0.196 | 0.190 | 3.487 | 0.174 | 0.159 | 3.558 | 26 | 0.147 | 0.143 |
| 12 | 0.966 | 0.733 | 0.759 | 0.527 | 0.432 | 0.410 | 6.329 | 0.059 | 0.057 | 6.510 | 29 | 0.047 | 0.048 |
| 16 | 1.324 | 1.015 | 1.046 | 0.762 | 0.638 | 0.603 | 9.256 | 0.027 | 0.033 | 9.124 | 23 | 0.024 | 0.025 |
| Swiss franc: | | | | | | | | | | | | | |
| 1 | 0.074 | 0.046 | 0.046 | 0.051 | 0.042 | 0.042 | 2.681 | 0.109 | 0.119 | 2.073 | 2 | 0.084 | 0.086 |
| 4 | 0.285 | 0.171 | 0.171 | 0.180 | 0.147 | 0.145 | 3.248 | 0.121 | 0.126 | 3.196 | 14 | 0.096 | 0.102 |
| 8 | 0.568 | 0.356 | 0.350 | 0.336 | 0.278 | 0.276 | 4.770 | 0.080 | 0.085 | 4.696 | 21 | 0.078 | 0.073 |
| 12 | 0.837 | 0.527 | 0.519 | 0.538 | 0.458 | 0.452 | 8.013 | 0.026 | 0.024 | 8.013 | 20 | 0.026 | 0.021 |
| 16 | 1.086 | 0.706 | 0.671 | 0.771 | 0.673 | 0.655 | 17.406 | 0.001 | 0.002 | 12.665 | 14 | 0.005 | 0.006 |
| Yen: | | | | | | | | | | | | | |
| 1 | 0.047 | 0.014 | 0.016 | 0.020 | 0.010 | 0.011 | 1.396 | 0.388 | 0.365 | 1.331 | 3 | 0.285 | 0.259 |
| 4 | 0.263 | 0.136 | 0.138 | 0.125 | 0.088 | 0.090 | 2.254 | 0.271 | 0.262 | 2.153 | 14 | 0.247 | 0.231 |
| 8 | 0.575 | 0.328 | 0.329 | 0.301 | 0.233 | 0.232 | 3.516 | 0.199 | 0.189 | 3.496 | 19 | 0.188 | 0.177 |
| 12 | 0.945 | 0.592 | 0.579 | 0.532 | 0.432 | 0.427 | 4.889 | 0.129 | 0.143 | 4.735 | 17 | 0.153 | 0.156 |
| 16 | 1.273 | 0.819 | 0.802 | 0.694 | 0.565 | 0.548 | 4.919 | 0.154 | 0.156 | 4.901 | 16 | 0.174 | 0.177 |

*Notes:* The table presents OLS estimates of the regression $e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,k}$, where $f_t \equiv (m_t - m_t^*) - (y_t - y_t^*)$. The (Gaussian) parametric and nonparametric bootstrap distributions are generated under the null hypothesis that the regressor follows an AR(4) for the Canadian dollar, the Swiss franc, and the yen, and an AR(5) for the deutsche mark. Exchange rates are dollars per unit of foreign currency. Adj-p and Adj-n are bias-adjusted values obtained by subtracting median values generated by the parametric and nonparametric bootstrap distributions, respectively, from the estimates. MSL-p and MSL-n are, respectively, the parametric and nonparametric bootstrap marginal significance levels for a one-tail test. $A$ is the truncation lag determined by Andrews's (1991) univariate AR(1) rule used for constructing the $t$ ratios with the data.

## B. Regression Estimates

Table 2 contains the estimated slope coefficients, $R^2$'s, asymptotic $t$ ratios, and aspects of their bootstrap distributions. At a given horizon, the slope coefficient is positive under the alternative hypothesis, and simple tests of the null hypothesis can be performed with one-tail tests on the $t$ ratios.

From columns (ii) and (v) of the table, one sees that for $k = 1$ the slope coefficients and the $R^2$'s for all four currencies are small in magnitude. Lengthening the forecast horizon, however, results in rising values of $\hat{\beta}_k$ and $R_k^2$ and rising values of the

asymptotic $t$ ratios for the deutsche mark, the Swiss franc, and the yen. In the case of the deutsche mark, at the 1-quarter horizon, $\hat{\beta}_1 = 0.04$, $R_1^2 = 0.02$, $t_1(20) = 1.84$, and $t_1(A) = 0.93$, while at the 16-quarter horizon, $\hat{\beta}_{16} = 1.32$, $R_{16}^2 = 0.76$, $t_1(20) = 9.26$, and $t_1(A) = 9.12$. This value of $\hat{\beta}_{16}$ implies that a 40-percent overvaluation of the dollar relative to the deutsche mark (which the estimates suggest was the case in 1985) predicts a depreciation of the dollar by 13.6 percent per annum over the next four years. The improved fit attained as $k$ increases suggests that the noise that dominates quarter-to-quarter changes in $e_t$ averages out over long horizons.
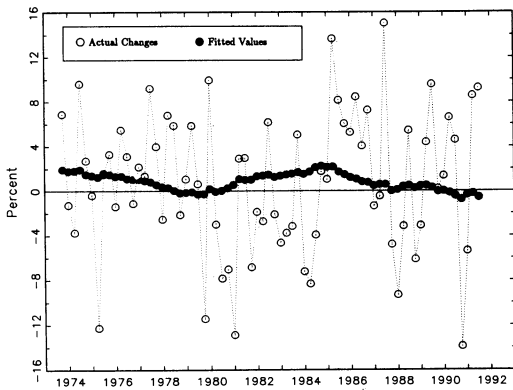
FIGURE 1. ONE-QUARTER CHANGES IN THE LOG
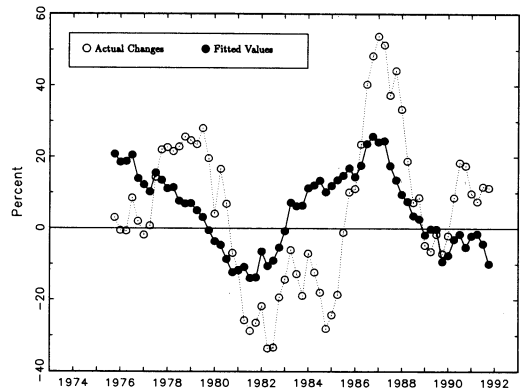DOLLAR/DEUTSCHE-MARK EXCHANGE RATE



FIGURE 3. EIGHT-QUARTER CHANGES IN THE
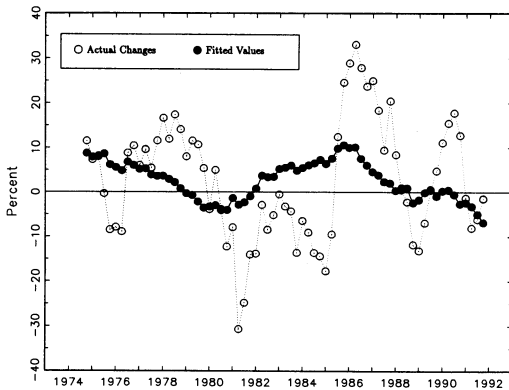LOG DOLLAR/DEUTSCHE-MARK EXCHANGE
RATE



FIGURE 2. FOUR-QUARTER CHANGES IN THE LOG
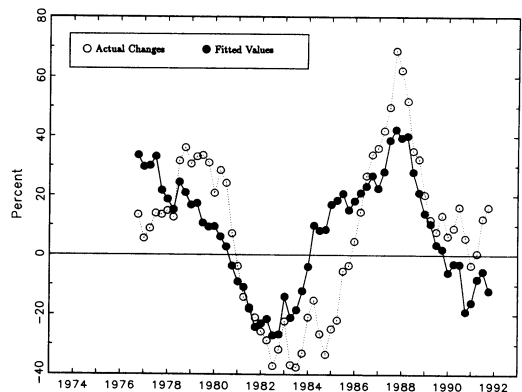DOLLAR/DEUTSCHE-MARK EXCHANGE RATE



FIGURE 4. TWELVE-QUARTER CHANGES IN THE
LOG DOLLAR/DEUTSCHE-MARK EXCHANGE
RATE

Figures 1 through 5 plot the actual and
fitted log exchange-rate changes for the
deutsche mark at each of the five horizons
(plots for the Swiss franc and the yen are
qualitatively similar and are suppressed to
save on space). The fitted values are indi-
cated by solid circles, and the actual
$k$-period log exchange-rate changes are in-
dicated by open circles. These figures illus-
trate the striking improvement in fit that
occurs as the forecast horizon is lengthened.

Columns (iii) and (iv) of Table 2 display
the bias-adjusted slope coefficients, and
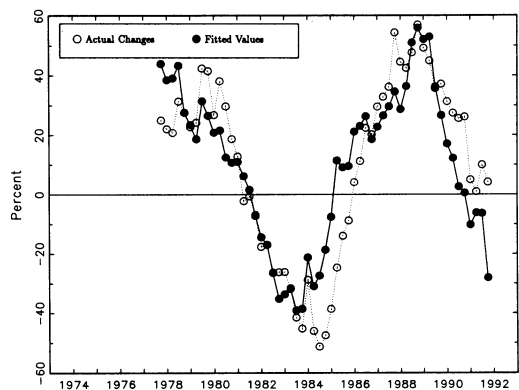columns (vi) and (vii) display the bias-



FIGURE 5. SIXTEEN-QUARTER CHANGES IN THE
LOG DOLLAR/DEUTSCHE-MARK EXCHANGE
RATE

adjusted $R^2$'s. As suggested in Subsection III-C, the bias estimated by the bootstrap median grows with the forecast horizon because extending the data overlap causes the effective sample size to shrink. The bias increases at a slower rate than the point estimates, however, so the adjusted values generally increase with the horizon as well. Such is the case for the deutsche mark. Under the parametric bootstrap, the bias-adjusted slope coefficient (Adj-p) increases from 0.012 at $k = 1$, to 1.015 at $k = 16$, while the adjusted $R^2$'s increase from 0.01 at $k = 1$ to 0.64 at $k = 16$. The slope coefficients adjusted by the nonparametric bootstrap median (Adj-n) range from 0.016 at $k = 1$ to 1.046 at $k = 16$, while the adjusted $R^2$'s range from 0.01 at $k = 1$ to 0.60 at $k = 16$.

Looking at columns (ix), (x), (xiii), and (xiv) in Table 2, it can be seen that the bootstrapped marginal significance levels of the asymptotic $t$'s generally shrink with the horizon for the deutsche mark, the Swiss franc, and the yen. The strongest evidence that an individual $t$ ratio is significantly positive comes at the 16-quarter horizon, where the marginal significance levels for one-tail tests under the nonparametric bootstrap for $t_{16}(A)$ are 0.03 for the deutsche mark, 0.01 for the Swiss franc, and 0.18 for the yen.[12]

A proper test of the null hypothesis involves testing whether the asymptotic $t$ ratios are *jointly* equal to zero across the five horizons for which the regressions are estimated. A natural way to proceed is to stack the $t$ ratios estimated from the data at each horizon into a vector and to use a quadratic measure of distance to test whether this vector is significantly different from the bootstrap median. However, Campbell (1992) has found that there are circumstances under which the power of the $t$ test against a fixed alternative increases with the forecast horizon, and in a related context, Pierre Perron (1989) and Shiller and Perron (1985) find that shortening the horizon lowers the power of tests for a random walk. Consequently, the low power inherent in testing whether the short-horizon $t$ ratio is zero can become impounded in the joint test statistic across all horizons.

Instead, I employ a procedure that may have more power against the alternative by testing whether the largest $t$ ratio among the five horizons is greater than zero.[13] Table 3 displays bias-adjusted values of $\max_k\{t_k(20)$: $k = 1, 4, 8, 12, 16\}$, $\max_k\{t_k(A)$: $k = 1, 4, 8, 12, 16\}$, and their bootstrapped marginal significance levels for a one-tail test. The joint test provides strong evidence against the null hypothesis for the deutsche mark and the Swiss franc. The evidence for the yen is less convincing, with marginal significance levels ranging from 0.25 to 0.26, while the evidence for the Canadian-dollar comes principally at the short horizons.

### C. Out-of-Sample Forecasts

I take $t_0 = T - 40$, which produces a sequence of 40 1-quarter horizon forecasts and 25 16-quarter horizon forecasts. Forecasting thus begins with 1981:4. The data reserved for the out-of-sample forecast experiment include an important turning point that occurred in 1985. The dollar experienced a large and sustained appreciation during the early 1980's, reaching a peak in 1985. From that peak, the dollar dropped sharply, losing nearly half of its 1985 value

---

[12] It is interesting to note that there is severe size distortion in the asymptotic $t$ tests since the inferences drawn from the bootstrapped $t$ distributions differ sharply from the large-sample inference. In addition, only a small part of the overstated asymptotic significance can be attributed to the bias in the slope-coefficient estimate. Looking at the deutsche mark regression at the 16-quarter horizon, the asymptotic standard error implied by $\hat{\beta}_{16} = 1.32$ and $t_{16}(A) = 9.12$ is 0.15. Dividing this number into the parametric bootstrap bias-adjusted slope-coefficient yields a $t$ ratio of 7.25, which still has an asymptotic tail probability of 0.

[13] I thank Pierre Perron for suggesting this test to me.

TABLE 3—JOINT TESTS

| Currency | Parametric bootstrap | | Nonparametric bootstrap | |
|---|---|---|---|---|
| | Bias-adjusted $\max_k\{t_k(A)\}$ | Bias-adjusted $\max_k\{t_k(20)\}$ | Bias-adjusted $\max_k\{t_k(A)\}$ | Bias-adjusted $\max_k\{t_k(20)\}$ |
| Canadian dollar | 0.703 | 1.120 | 0.594 | 0.969 |
| | (0.384) | (0.322) | (0.408) | (0.348) |
| Deutsche mark | 6.918 | 6.942 | 6.761 | 7.064 |
| | (0.036) | (0.046) | (0.049) | (0.052) |
| Swiss franc | 9.699 | 14.903 | 9.592 | 14.820 |
| | (0.010) | (0.001) | (0.009) | (0.005) |
| Yen | 2.091 | 2.181 | 2.031 | 2.078 |
| | (0.267) | (0.248) | (0.273) | (0.256) |

*Notes:* Bias is estimated by the bootstrap median. Numbers in parentheses are the marginal significance levels generated by the parametric or nonparametric bootstrap distribution for the one-tail test that $\max_k\{t_k(A)\} > 0$ or $\max_k\{t_k(20)\} > 0$.

by 1987, and it has remained relatively stable since that time.

The postsample prediction performance of the regression, the driftless random walk, and the in-sample regression fit are displayed in Table 4. To facilitate a comparison among the models, I report relative RMSE performance in columns (ii)–(iv). Several features of this table are worth pointing out. As expected, there is some deterioration in the out-of-sample forecast's precision relative to the in-sample fit. Column (iii), labeled OUT/IN, indicates that the deterioration is seen to be most severe for the Canadian dollar: at the 12- and 16-quarter horizons, the out-of-sample forecast's RMSE is more than double the size of the in-sample fit's RMSE. The deterioration for the other three currencies is modest by comparison. At the 16-quarter horizon, the RMSE of the out-of-sample prediction exceeds that of the in-sample fit by only 20 percent for the deutsche mark and the Swiss franc and by 60 percent for the yen.

The regression forecast RMSE relative to that of the driftless random walk is displayed in column (iv). Here, it can be seen that the regression forecasts beat the random walk at every horizon for the Swiss franc and the yen, and at the 12- and 16-quarter horizons for the deutsche mark. At the shorter horizons, the improvement over the random walk is small, but at the 16-quarter horizon, the regression's out-of-sample RMSE is roughly half that of the driftless random walk for the deutsche mark, the Swiss franc, and the yen. The marginal significance levels for (OUT/RW) are also seen to decline with $k$ for these currencies. Consistent with the results from Table 2, however, the Canadian-dollar regression outperforms the random walk only at the 1-quarter horizon.

The $\mathcal{DM}$ statistics, displayed in columns (vii) and (x), are set up to be positive when the regression outperforms the random walk. Observe that their bootstrapped marginal significance levels generally decline with $k$, which again suggests that the strongest evidence against the null hypothesis comes at the longer forecast horizons. There is substantial size distortion at the longer horizons as well. For example, the nonparametric bootstrap marginal significance level for $\mathcal{DM}(A) = 8.7$ is 0.01 for the deutsche mark at the 16-quarter horizon. However, unlike the asymptotic $t$ ratios, the $\mathcal{DM}$ statistics are sensitive to the method by which the truncation lag of the Bartlett window is chosen, which suggests that the sampling properties of the (OUT/RW) statistic may be more reliable than those of $\mathcal{DM}$ at the longer horizons.

The joint tests are again conducted using the distribution of the statistic's extreme value over the five horizons. The extreme values are the maximum values of $\mathcal{DM}$ and the minimum values of the (OUT/RW)

TABLE 4—OUT-OF-SAMPLE FORECAST EVALUATION

| (i) k | (ii) IN/RW | (iii) OUT/IN | (iv) OUT/RW | (v) MSL-p | (vi) MSL-n | (vii) $\mathcal{DM}(20)$ | (viii) MSL-p | (ix) MSL-n | (x) $\mathcal{DM}(A)$ | (xi) A | (xii) MSL-p | (xiii) MSL-n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Canadian dollar: | | | | | | | | | | | | |
| 1 | 0.960 | 1.040 | 0.998 | 0.209 | 0.194 | 0.061 | 0.215 | 0.202 | 0.036 | 1 | 0.218 | 0.201 |
| 4 | 0.889 | 1.258 | 1.119 | 0.571 | 0.538 | −1.270 | 0.526 | 0.487 | −0.925 | 8 | 0.494 | 0.468 |
| 8 | 0.675 | 1.695 | 1.145 | 0.447 | 0.397 | −1.036 | 0.427 | 0.377 | −0.890 | 17 | 0.420 | 0.390 |
| 12 | 0.654 | 2.197 | 1.436 | 0.613 | 0.578 | −1.916 | 0.574 | 0.556 | −1.661 | 18 | 0.587 | 0.579 |
| 16 | 0.799 | 2.128 | 1.699 | 0.654 | 0.636 | −2.596 | 0.578 | 0.542 | −1.857 | 15 | 0.567 | 0.555 |
| Deutsche mark: | | | | | | | | | | | | |
| 1 | 0.988 | 1.027 | 1.015 | 0.397 | 0.339 | −0.932 | 0.458 | 0.393 | −0.846 | 4 | 0.536 | 0.493 |
| 4 | 0.927 | 1.120 | 1.037 | 0.345 | 0.288 | −1.345 | 0.563 | 0.511 | −0.852 | 9 | 0.478 | 0.427 |
| 8 | 0.833 | 1.203 | 1.002 | 0.268 | 0.217 | −0.027 | 0.270 | 0.220 | −0.020 | 18 | 0.270 | 0.221 |
| 12 | 0.670 | 1.188 | 0.796 | 0.127 | 0.092 | 4.246 | 0.068 | 0.059 | 0.094 | 16 | 0.151 | 0.136 |
| 16 | 0.431 | 1.216 | 0.524 | 0.040 | 0.025 | 8.719[a] | 0.061 | 0.047 | 8.719 | 18 | 0.021 | 0.011 |
| Swiss franc: | | | | | | | | | | | | |
| 1 | 0.972 | 1.026 | 0.997 | 0.305 | 0.266 | 0.066 | 0.320 | 0.278 | 0.064 | 3 | 0.315 | 0.271 |
| 4 | 0.886 | 1.108 | 0.981 | 0.291 | 0.263 | 0.218 | 0.304 | 0.272 | 0.162 | 12 | 0.298 | 0.274 |
| 8 | 0.780 | 1.176 | 0.917 | 0.256 | 0.219 | 0.703 | 0.260 | 0.236 | 0.560 | 17 | 0.253 | 0.227 |
| 12 | 0.625 | 1.181 | 0.738 | 0.152 | 0.132 | 2.933 | 0.161 | 0.137 | 0.938 | 13 | 0.255 | 0.211 |
| 16 | 0.335 | 1.229 | 0.411 | 0.033 | 0.023 | 9.650[b] | 0.080 | 0.058 | 1.996 | 8 | 0.192 | 0.159 |
| Yen: | | | | | | | | | | | | |
| 1 | 0.962 | 1.027 | 0.988 | 0.304 | 0.257 | 1.571 | 0.168 | 0.132 | 0.836 | 3 | 0.177 | 0.134 |
| 4 | 0.822 | 1.129 | 0.928 | 0.257 | 0.207 | 2.302 | 0.151 | 0.118 | 1.487 | 10 | 0.134 | 0.105 |
| 8 | 0.688 | 1.191 | 0.819 | 0.214 | 0.162 | 3.096 | 0.142 | 0.117 | 1.803 | 13 | 0.152 | 0.117 |
| 12 | 0.536 | 1.329 | 0.712 | 0.196 | 0.148 | 3.319 | 0.174 | 0.148 | 1.147 | 17 | 0.164 | 0.135 |
| 16 | 0.363 | 1.579 | 0.574 | 0.152 | 0.119 | 5.126 | 0.178 | 0.160 | 3.096 | 16 | 0.151 | 0.131 |

*Notes:* The table presents ratios of root-mean-squared errors for the regression's out-of-sample forecasts (OUT), the driftless random walk (RW), and the in-sample regression residual during the forecast period (IN). The first forecast is made on 1981:4. $\mathcal{DM}(20)$ and $\mathcal{DM}(A)$ are the Diebold-Mariano statistics constructed using the method of Newey and West (1987) with the truncation lag of the Bartlett window set to 20 and set by Andrews's (1991) AR(1) rule, respectively. In instances where the estimated spectral density at frequency zero of the squared error differential is nonpositive (see footnote 8), the Bartlett-window truncation lag is decreased by 1. MSL-p and MSL-n are marginal significance levels, generated by the parametric and nonparametric bootstrap distributions, respectively, for one-tail tests.
[a] Bartlett-window truncation lag = 18.
[b] Bartlett-window truncation lag = 17.

RMSE ratio across the five horizons. Table 5 displays bias-adjusted values of the extreme values and bootstrapped marginal significance levels. Here, the null hypothesis can be rejected for the deutsche mark with $\mathcal{DM}(A)$ at the 5-percent level (MSL-p = 0.03, MSL-n = 0.02), and for the Swiss franc using $\mathcal{DM}(20)$ near the 10-percent level (MSL-p = 0.10, MSL-n = 0.08). The evidence against the null hypothesis for the yen is less forcible [MSL-n = 0.16 for $\mathcal{DM}(A)$]. Finally, using the bootstrapped (OUT/RW) RMSE ratios, the null hypothesis is rejected at the 5-percent level for both the deutsche mark (MSL-p =

0.042, MSL-n = 0.026) and the Swiss franc (MSL-p = 0.04, MSL-n = 0.02), with slightly weaker evidence for the yen (MSL-p = 0.16 and MSL-n = 0.13).

## V. Concluding Remarks

This paper has presented evidence that there is an economically significant predictable component in long-horizon changes in log exchange rates. The evidence comes from regressions of long-horizon changes in log exchange rates on the current log exchange rate's deviation from a linear combination of log relative money stocks and log

TABLE 5—OUT-OF-SAMPLE FORECASTS, JOINT TESTS

| Currency | $\min_k\{(OUT/RW)_k\}$ | | $\max_k\{\mathcal{DM}_k(20)\}$ | | $\max_k\{\mathcal{DM}_k(A)\}$ | |
|---|---|---|---|---|---|---|
| | Adj-p (MSL-p) | Adj-n (MSL-n) | Adj-p (MSL-p) | Adj-n (MSL-n) | Adj-p (MSL-p) | Adj-n (MSL-n) |
| Canadian dollar | 0.986 (0.425) | 0.983 (0.404) | 0.431 (0.420) | 0.585 (0.404) | 0.312 (0.423) | 0.391 (0.404) |
| Deutsche mark | 0.512 (0.042) | 0.505 (0.026) | 9.119 (0.081) | 9.320 (0.064) | 8.991 (0.030) | 9.139 (0.019) |
| Swiss franc | 0.433 (0.035) | 0.416 (0.024) | 9.329 (0.102) | 9.549 (0.082) | 1.771 (0.241) | 1.924 (0.206) |
| Yen | 0.646 (0.157) | 0.620 (0.126) | 4.222 (0.222) | 4.443 (0.207) | 2.501 (0.188) | 2.621 (0.161) |

*Notes:* The first forecast is made at 1981:4. The minima (maxima) are taken over the five horizons. Bias is estimated by the bootstrap median. Bias-adjusted values using the parametric (nonparametric) bootstrap distribution are denoted by Adj-p (Adj-n). MSL-p (MSL-n) is the marginal significance level generated by the parametric (nonparametric) bootstrap distribution for the one-tail test that $\min_k(OUT/RW)_k = 1$ or $\max_k\{\mathcal{DM}_k\} = 0$.

relative real income. While short-horizon changes tend to be dominated by noise, this noise is apparently averaged out over time, thus revealing systematic exchange-rate movements that are determined by economic fundamentals. These findings are noteworthy because it has long been thought that log exchange rates are unpredictable.

The coefficient estimates display a pattern suggested by models with nominal rigidities that cause the exchange rate to adjust gradually in response to either nominal or real shocks. In particular, the bias-adjusted regression slope coefficients and $R^2$'s increase with the forecast horizon. For three of the four exchange rates studied, the out-of-sample forecasts of the regression outperform those of the driftless random walk at the longer horizons. The forecastability of long-horizon changes may be a common feature in economic time series. Campbell and Shiller (1988) and Fama and French (1988) have found multiple-period stock returns to be predictable, and Frederic S. Mishkin (1990a,b) has found long-horizon inflation to be predictable as well.

The available time series is relatively short, however, which makes asymptotic inference unreliable. I account for problems of small-sample bias and slow convergence to the asymptotic distribution by drawing inference from bootstrap distributions generated under the null hypothesis of exchange-rate unpredictability. While the statistical procedures employed are unable to provide a blanket rejection of the null hypothesis at the standard 5-percent level, it is important to emphasize that the reason for this may simply be that the time series from the float is not sufficiently long. Recall that the strongest evidence against the null hypothesis comes at the horizon for which there are the fewest observations. At the 16-quarter horizon, the in-sample results are based on fewer than five nonoverlapping observations. While the effective number of observations is undoubtedly greater than 5, it is probably closer to 5 than to 76. Similarly, the out-of-sample forecasts at the 16-quarter horizon amount to fewer than two nonoverlapping observations. Because of the extremely small sample available, the significance levels in tests of the null hypothesis should be adjusted up from 5 percent. Upon taking the sample size into consideration, the statistical evidence from the bootstrap distributions weighs against the hypothesis of equal forecast ability.

The goal of the paper was to investigate the extent to which deviations of the exchange rate from a fundamental value suggested by economic theory are useful in predicting exchange-rate changes over long

horizons. The empirical work was purposely restricted to an analysis of a regression on one variable so that the predictive relation could be conveniently characterized. Clearly, a number of complications can be introduced to the empirical analysis, and such modification may result in improved forecast accuracy. For example, the deviations from the fundamental value could be estimated from the data in a cointegrating regression, and the definition of the economic fundamentals might contain an expanded list of economic variables.

## REFERENCES

**Abauf, Niso and Jorion, Philippe.** "Purchasing Power Parity in the Long Run." *Journal of Finance*, March 1990, *45*(1), pp. 157–74.

**Andrews, Donald W. K.** "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica*, May 1991, *59*(3), pp. 817–58.

**Baillie, Richard T. and Bollerslev, Tim.** "The Message in Daily Exchange Rates: A Conditional Variance Tale." *Journal of Business and Economic Statistics*, July 1989, *7*(3), pp. 1297–1305.

**Baxter, Marianne.** "Real Exchange Rates and Real Interest Differentials: Have We Missed the Business-Cycle Relationship?" Mimeo, University of Virginia, 1993.

**Bilson, John F. O.** "Rational Expectations and the Exchange Rate," in Jacob A. Frenkel and Harry G. Johnson, eds., *The economics of exchange rates: Selected studies*. Reading, MA: Addison-Wesley, 1978, pp. 75–96.

**Blough, Stephen R.** "The Relationship between Power and Level for Generic Unit Root Tests in Finite Samples." *Journal of Applied Econometrics*, July-September 1992, *7*(3), pp. 295–308.

**Campbell, John Y.** "Why Long Horizons? Power Against Persistent Alternatives." Mimeo, Princeton University, 1992.

**Campbell, John Y. and Shiller, Robert J.** "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance*, July 1988, *43*(3), pp. 661–76.

**Cumby, Robert E. and Huizinga, John.** "The

Predictability of Real Exchange Rate Changes in the Short and Long Run." *Japan and the World Economy*, 1991, *3*(1), pp. 17–38.

**Diebold, Francis X.; Husted, Steven and Rush, Mark.** "Real Exchange Rates Under the Gold Standard." *Journal of Political Economy*, December 1991, *99*(6), pp. 1252–71.

**Diebold, Francis X. and Mariano, Roberto S.** "Comparing Predictive Accuracy." Mimeo, University of Pennsylvania, 1993.

**Diebold, Francis X. and Nason, James A.** "Nonparametric Exchange Rate Prediction?" *Journal of International Economics*, May 1990, *28*(3-4), pp. 315–32.

**Dornbusch, Rudiger.** "Expectations and Exchange Rate Dynamics." *Journal of Political Economy*, December 1976, *84*(6), pp. 1161–76.

**Driskill, Robert A.; Mark, Nelson C. and Sheffrin, Steven M.** "Some Evidence in Favor of a Monetary Rational Expectations Exchange Rate Model with Imperfect Capital Substitutability." *International Economic Review*, February 1992, *33*(1), pp. 223–37.

**Edison, Hali J.** "Purchasing Power Parity in the Long Run: A Test of the Dollar/ Pound Exchange Rate (1890–1978)." *Journal of Money, Credit, and Banking*, August 1987, *19*(3), pp. 376–87.

**Engel, Charles and Hamilton, James D.** "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review*, September 1990, *80*(4), pp. 689–713.

**Engle, Robert F.** "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica*, July 1982, *50*(4), pp. 987–1007.

**Fama, Eugene F. and French, Kenneth R.** "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, October 1988, *22*(1), pp. 3–25.

**Fisher, Eric O. and Park, Joon Y.** "Testing Purchasing Power Parity Under the Null Hypothesis of Co-integration." *Economic Journal*, November 1991, *101*(409), pp. 1476–84.

**Frankel, Jeffrey A.** "On the Mark: A Theory

of Floating Exchange Rates Based on Real Interest Differentials." *American Economic Review*, September 1979, *69*(4), pp. 610–22.

Frenkel, Jacob A. "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence." *Scandinavian Journal of Economics*, May 1976, *78*(2), pp. 200–24.

Grilli, Vittorio and Kaminsky, Graciela. "Nominal Exchange Rate Regimes and the Real Exchange Rate: Evidence from the United States and Great Britain, 1885–1986." *Journal of Monetary Economics*, April 1991, *27*(2), pp. 191–212.

Hansen, Lars P. and Hodrick, Robert J. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *Journal of Political Economy*, October 1980, *88*(5), pp. 829–53.

Hodrick, Robert J. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 1992, *5*(3), pp. 357–86.

Huizinga, John. "An Empirical Investigation of the Long Run Behavior of Real Exchange Rates." *Carnegie-Rochester Conference Series on Public Policy*, Autumn 1987, *27*, pp. 149–214.

Kim, Yoonbai. "Purchasing Power Parity in the Long Run: A Cointegration Approach." *Journal of Money, Credit, and Banking*, November 1990, *22*(4), pp. 491–503.

Krugman, Paul. "Target Zones and Exchange Rate Dynamics." *Quarterly Journal of Economics*, August 1991, *106*(3), pp. 311–25.

*Main Economic Indicators*. Paris: Organization for Economic Cooperation and Development, 1992.

Meese, Richard. "Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates?" *Journal of Political Economy*, April 1986, *94*(2), pp. 345–73.

Meese, Richard and Rogoff, Kenneth. "Empirical Exchange Rate Models of the 1970's: Do They Fit Out of Sample?" *Journal of International Economics*, February 1983a, *14*(1/2), pp. 3–24.

_____. "The Out-of-Sample Failure of Exchange Rate Models: Sampling Error or Misspecification?" in Jacob Frenkel, ed., *Exchange rates and international macroeconomics*. Chicago: University of Chicago Press, 1983b, pp. 67–105.

_____. "Was It Real? The Exchange-Rate Interest Differential Relation Over the Modern Floating-Rate Period." *Journal of Finance*, September 1988, *43*(4), pp. 933–48.

Meese, Richard and Singleton, Kenneth J. "Rational Expectations and the Volatility of Floating Exchange Rates." *International Economic Review*, October 1983, *24*(3), pp. 721–73.

Mishkin, Frederic S. "What Does the Term Structure Tell Us About Future Inflation." *Journal of Monetary Economics*, January 1990a, *25*(1), pp. 77–95.

_____. "The Information in the Longer Maturity Term Structure About Future Inflation." *Quarterly Journal of Economics*, August 1990b, *105*(3), pp. 815–28.

Mussa, Michael. "The Exchange Rate, the Balance of Payments, and Monetary and Fiscal Policy Under a Regime of Controlled Floating." *Scandinavian Journal of Economics*, May 1976, *78*(2), pp. 229–48.

_____. "A Model of Exchange Rate Dynamics." *Journal of Political Economy*, February 1982, *90*(1), pp. 74–104.

Nelson, Charles R. and Kim, Myung J. "Predictable Stock Returns: The Role of Small Sample Bias." *Journal of Finance*, June 1993, *48*(2), pp. 641–61.

Newey, Whitney K. and West, Kenneth D. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, May 1987, *55*(3), pp. 703–8.

Perron, Pierre. "Testing for a Random Walk: A Simulation Experiment of Power When the Sampling Interval Is Varied," in B. Raj, ed., *Advances in econometrics and modeling*. Dordrecht: Kluwer, 1989, pp. 47–68.

Rush, Mark and Husted, Steven. "Purchasing Power Parity in the Long Run." *Canadian Journal of Economics*, February 1985, *18*(1), pp. 137–45.

Shiller, Robert J. and Perron, Pierre. "Testing

the Random Walk Hypothesis: Power versus Frequency of Observation." *Economics Letters*, 1985, *18*(4), pp. 381–86.

**Stambaugh, Robert F.** "Bias in Regression with Lagged Stochastic Regressors." Mimeo, University of Chicago, 1986.

**Stock, James H.** "Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series." *Journal of Monetary Economics*, December 1991, *28*(3), pp. 435–60.

**West, Kenneth D.** "A Standard Monetary Model and the Variability of the Deutschemark–Dollar Exchange Rate." *Journal of International Economics*, August 1987, *23*(1/2), pp. 57–76.

**Woo, Wing T.** "The Monetary Approach to Exchange Rate Determination Under Rational Expectations: The Dollar–Deutschmark Rate." *Journal of International Economics*, February 1985, *18*(1/2), pp. 1–16.