

Bootstrap Predictability of Daily Exchange Rates¹

Demosthenes N. Tambakis
City University Business School

Anne-Sophie Van Royen
FERC, City University Business School
University College London

Revised January 14, 1999

Abstract

We develop a measure of time series predictability and apply it to evaluate alternative parametric models' out-of-sample performance. Predictability is a nonlinear statistic of a forecasting model's relative expected losses at two different forecast horizons which allows flexible choice of both the estimation and loss function specifications. AR-GARCH models are fitted to daily exchange rate returns, and bootstrap inference is used to assess the data's predictability under mean squared error (MSE) loss. We compute daily predictability point estimates and confidence intervals for the dollar exchange rate returns of the deutschemark, pound sterling and the yen. By comparing the data's predictability using the fitted models to that using a random walk under MSE loss, we find that all three exchange rate returns are less predictable using AR-GARCH. The results also suggest that the pound's actual returns are relatively more predictable than the other time series.

Keywords: Predictability, bootstrap, loss function, exchange rates

JEL Classification: C14, C53, F31

¹We would like to thank Alex Michaelides, Mark Salmon and conference participants at Computational Finance 99 in New York for helpful comments and discussions. All errors remain our own. Corresponding author's address: D.N. Tambakis, City University Business School, Frobisher Crescent, The Barbican, London EC2Y 8HB, U.K. Tel: 44-171-477-8761, Fax: 44-171-477-8648. E-mail address: d.tambakis@city.ac.uk.

1 Introduction

At what forecast horizon is one time series more predictable than another using a given model? Alternatively, how much more (or less) predictable is a given time series when modeled using an estimated model rather than a random walk? These questions have attracted little attention despite the importance attached to assessing the forecast accuracy of different models. Recently, researchers are acknowledging that data predictability cannot be defined independently of the forecasting model used, as well as the loss function employed for forecast evaluation. However, conditional upon a particular model and loss function specification, data predictability can be assessed at various horizons of interest.² Therefore, the need arises for procedures estimating predictability so that comparisons of it can be made across datasets, as well as different estimation procedures and forecast loss functions.

The k -step-ahead predictability of covariance-stationary time series was defined by Granger and Newbold (1986) and Beran (1994) as a relative measure of forecast accuracy: it is the proportion of the unobservable unconditional variance explained by the mean squared error of a k -step ahead conditional forecast. Diebold and Kilian (1997) generalized this notion and defined relative predictability as inversely related to the ratio of the expected losses of conditionally optimal linear forecasts at short and long forecast horizons.

In contrast to Granger and Newbold's absolute predictability definition, the Diebold-Kilian measure has the advantage that the benchmark level of forecast accuracy which is an input to the process is directly observable. The choice of horizons is flexible, with the long horizon corresponding to the benchmark level of forecast accuracy. Moreover, the loss function used for forecast evaluation can be quite general and determined by the application at hand.³ Diebold-Kilian computed predictability confidence intervals by fitting AR models to U.S. macro data and evaluating expected losses according to mean squared error loss. Because of the nonlinearity of the predictability

²See Diebold and Kilian (1997), Diebold (1998) and Rothenberg and Woodford (1996).

³The definition is also asymptotically valid for covariance nonstationary time series.

statistic, bootstrap methods were used to obtain the confidence intervals.⁴

This paper extends the Diebold-Kilian framework to evaluate the forecasting performance of alternative parametric exchange rate models according to MSE. We compute bias-corrected predictability confidence intervals for daily returns of the British pound, Japanese yen and German mark's exchange rates against the US dollar. Our motivation for using the predictability statistic is the failure of most parametric and nonparametric models of exchange rate returns to outperform the random walk at high frequencies.⁵ Because it is defined as a ratio, predictability is a relative statistic, circumventing the stylized failure of forecasting models by comparing the evolution of their out-of-sample expected loss.

The (j, k) -step-ahead predictability $P(L, \Omega, j, k)$ of exchange rate returns is computed by fixing the long forecast horizon k to one month-ahead (22 days) and varying the short forecast horizon j from 1 to 21 days. This is done by means of the following procedures. First, an AR-GARCH model is fitted to in-sample returns data ($N - 22$) using standard Box-Jenkins methodology. The mean equation follows a simple autoregressive process, and the conditional variance equation is modeled using a GARCH(1,1) specification. The chosen model for each exchange rate is then used to generate conditionally optimal linear forecasts of the true data for the out-of-sample range: $1 \leq j \leq 21, k = 22$. These are the "true forecasts" from the selected model. By comparing the true forecasts to the last 22 out-of-sample observations we compute the "true", or *baseline*, value of the predictability statistic. The latter is required for the construction of bias-corrected predictability confidence intervals.

⁴The bootstrap has the advantage of being nonparametric; moreover, bootstrap statistics can be estimated consistently using least squares, provided the model residuals used for resampling are iid. Detailed treatments of bootstrapping include Efron and Tibshirani, (1986,1993), Hall (1992) and Hall and Horowitz (1996).

⁵On structural models's worse out-of-sample performance against a random walk using MSE see Meese and Rogoff (1983) and Berkowitz and Giorgianni (1997). For time series models see Brooks (1997), Diebold and Nason (1990) and Satchell and Timmermann (1995).

In the second procedure, we generate 500 bootstrap replications of the true data, using block resampling with replacement from the "true" errors from the fitted model. We then evaluate the predictability statistic for each replication under MSE loss and compute the 95% bias-corrected (BC) confidence intervals. After resampling $b = 1, \dots, B$ times, each resampled error vector is used to generate a pseudo-data vector by applying it to the estimated coefficients for the true data. Estimation is carried out on each pseudo-data vector, using the same lag structure as for the true data. Each of the resulting B pseudo-models can then be used to construct pseudo-forecasts of the out-of-sample period of length 22. The value of $\hat{P}_b(j, k)$ for each bootstrap replication is obtained by evaluating its pseudo-forecast vector against the corresponding pseudo-data vector at all short (j) forecast horizons according to MSE. Finally, we compute the $\hat{P}_b(j, k)$'s for all B bootstrap replications and obtain mean predictability point estimates and bias-corrected (BC) bootstrap confidence intervals.

There are two main findings. First, all 3 exchange rate returns are more predictable under MSE loss when using a random walk than when using the chosen AR-GARCH model for all intra-month forecast horizons. Such a comparison is feasible because forecasting returns using a random walk under MSE display a linearly declining predictability pattern. This can be directly compared to that obtained when a different model is used to forecast the data, thus making the random walk a convenient benchmark of forecasting models' out-of-sample performance.

Second, examining the variation of returns predictability with the forecast horizon, we find that £/\$ exchange rate returns have the higher predictability across intra-month forecast horizons, while the DM/\$ return has the lower predictability. However, the 95% confidence intervals for the predictability of DM/\$ and yen/\$ returns are both narrower than that of £/\$. Given the explanatory power of the respective time series models, the economic intuition for the relatively higher predictability of £/\$ returns may have to do with the historical "coupled" relationship between the pound and the dollar.

The remainder of the paper is arranged as follows. In section 2 relative predictability is defined and its key properties reviewed. The exchange rate

data along with their statistical properties and estimated models are presented in section 3. Section 4 develops the bootstrap technique for computing relative predictability and discusses bias-corrected confidence intervals. Section 5 presents the results and their implications for daily exchange rate forecasting under mean square error loss. Section 6 concludes.

2 A review of predictability measures

Granger and Newbold (1986) and Beran (1994) define the predictability of univariate covariance stationary processes under MSE loss by analogy to the OLS formula for R^2 :

$$G(j, k) = \frac{\text{var}(\hat{y}_{t+j,t})}{\text{var}(y_{t+j})} = 1 - \frac{\text{var}(e_{t+j,t})}{\text{var}(y_{t+j})} \quad (1)$$

where $\hat{y}_{t+j,t}$ is the optimal linear (conditional mean) j -step-ahead forecast of y_t and $e_{t+j,t} = y_{t+j} - \hat{y}_{t+j,t}$ is the associated forecast error. $G(j, k)$ may be thought of as an *absolute* predictability measure, to the extent that the unconditional variance in the denominator is unobservable.

Diebold and Kilian (1997)—henceforth *DK*—extend this definition to covariance nonstationary and possibly multivariate time series processes under general loss functions. In line with the intuition of Granger and Newbold, the predictability of process y_t is inversely related to the ratio of the conditionally expected loss of an optimal short-run forecast, $E(L(e_{t+j,t}))$, to that of an optimal long-run forecast, $E(L(e_{t+k,t}))$, where $j < k$. Intuitively, if $E(L(e_{t+j,t})) < E(L(e_{t+k,t}))$ then the process is more predictable at horizon j relative to horizon k . In contrast, if $E(L(e_{t+j,t})) \simeq E(L(e_{t+k,t}))$ the process is almost equally predictable at horizon j relative to k , i.e. it is relatively unpredictable. DK thus define the predictability $P_t(j, k, L, \Omega)$ ($j < k$) of y_t to be:

$$P_t(L, \Omega_t, j, k) = 1 - \frac{E_t(L(e_{t+j,t}))}{E_t(L(e_{t+k,t}))} \quad (2)$$

where the information set Ω_t can be either univariate or multivariate and y_t can be either stationary or nonstationary. The choice of horizons j and k is flexible provided $k < \infty$.

An equivalent form of (2) uses the ratio of the expected forecast losses for period t made in periods $t - j$ and $t - k$ in the past:

$$P_t(L, \Omega_t, j, k) = 1 - \frac{E_{t-j}(L(e_{t,t-j}))}{E_{t-k}(L(e_{t,t-k}))} \quad (3)$$

Equation (3) relates the accuracy of a few steps ahead (j) vs. more steps ahead (k) forecasts of y_t to the parameters influencing predictability. In the context of univariate AR(1) processes, $y_t = \rho_1 y_{t-1} + u_t$, the j -step-ahead predictability of a white noise process ($\rho_1 = 0$) will be zero for all j , as short-run and long-run forecasts are equal. In contrast, the relative predictability of a random walk process ($\rho_1 = 1$) is linearly declining in the short-run forecast horizon: $P_t(j, k) = 1 - j/k$. Relations (2)-(3) imply that the bounds of the P_t statistic for any time series are $P_t(L, \Omega_t, j, k) \in (-\infty, 1]$, with larger values indicating higher relative predictability. For covariance stationary data, comparing predictability at two short term horizons j_1 and j_2 with $j_1 < j_2$, the information set used for j_2 -steps-ahead forecasts is likely to be poorer than that j_1 -steps-ahead, so we expect predictability to decline as we compare forecasts further into the future.⁶

In this paper we focus on univariate information sets and quadratic (mean squared error) loss.⁷ For univariate information set Ω_t under MSE loss we may write $P_t(L, \Omega_t, j, k) = P_t(j, k)$. Let N is the total number of observations, of which $N - k$ are reserved for estimation and k for out-of-sample forecast evaluation. The expected mean square error of j steps-ahead forecasts then is:

$$E_t(L(e_{t+j,t})) = MSE_j = \frac{1}{j} \sum_{i=N-k+1}^{N-k+j} (y_{t+i} - \hat{y}_{t+i,t})^2 \quad (4)$$

where the value of the time series at time t is y_t , and the conditionally optimal i -step ahead forecast of y made in period t is $\hat{y}_{t+i,t} = E_t(y_{t+i} | \Omega_t)$. Finally,

⁶For covariance stationary processes it can be shown that $0 < P_t(L, \Omega, j, k) < 1$, reflecting increased uncertainty over more distant forecasts. However, the predictability of covariance nonstationary series can also be negative, reflecting the fact that short-term expected losses may exceed the long-term ones.

⁷In general, however, the loss function need not be restricted to the quadratic specification. For the theory and applications of general—nonquadratic and/or asymmetric—loss functions see Diebold and Mariano (1995) and Christoffersen and Diebold (1996, 1997).

the Wold representation of covariance stationary series y_t is:

$$y_t = \sum_{i=0}^{\infty} c_i \varepsilon_{t-i} , \quad c_0 = 1$$

implying that the n step-ahead conditional forecast variance—the n -steps-ahead MSE loss—is increasing in n :

$$\frac{\partial}{\partial n} E_t(\hat{y}_{t,t+n} - y_{t+n})^2 > 0 ,$$

as all covariance terms are zero because ε is iid white noise. Therefore, fixing the long horizon at k and increasing the short term $j = n$ lowers relative predictability $P_t(j, k)$ under MSE, in line with the intuitive notion of forecast uncertainty increasing over time for covariance stationary series.⁸

3 The data

The univariate information set Ω_t for computing $P_t(L(\cdot), \Omega_t, j, k)$ consists of daily spot exchange rates for the British pound (GBP), German mark (DM) and Japanese yen (JPY) against the US dollar for the period from January 1, 1988 to April 7, 1998. Transforming spots into log returns to account for nonstationarity in the levels yields $N = 2,678$ observations. Tables 1-3 present summary statistics for exchange rate returns and the properties of the fitted models. Panels A summarize the distributional properties of returns. All ADF tests indicate stationarity. The in-sample returns variance is much larger than the mean, and their skewness is close to zero. Returns are strongly leptokurtic, thus strongly rejecting normality. The Ljung-Box Q_x statistic is not significant at 20 lags for JPY and DM, but it is significant for GBP. Panels B report the selected AR-GARCH model fitted to each returns time series. For each model, the autoregressive lag order is determined using the Akaike and Schwartz information criteria (AIC and SIC). SIC penalizes an overspecified lag structure, so it is less likely to overestimate the required lags than the AIC. However, selecting a larger number of lags yields more variable forecasts, i.e. ones which take more time to converge to the conditional mean.

⁸This is not true of any loss functional form. For example, under mean absolute percentage error (MAPE) loss, increasing the short horizon need not imply a lower $P(j, k)$.

The chosen AR orders thus vary from 5 to 11. The explanatory power is very low, peaking at just 0.01 for BP. A GARCH(1,1) specification is always preferred with no MA structure in the errors for GBP/\$ and DM/\$, and an ARMA(1,1) structure for the errors of JPY/\$ returns.

The standard set of diagnostics suggest that the errors from the GARCH models are uncorrelated both in the levels and squares. However, there is strong evidence of nonlinear dependence in daily exchange rate returns working through the conditional mean and higher moments, hence the iid requirement is rejected.⁹ This implies that estimation procedures have residuals which are not iid, thus yielding inconsistent parameter estimates. The issue of bootstrap resampling from the true errors using a block index thus becomes important. As bootstrap replications are drawn with replacement from an empirical distribution of true errors which in general are not iid, we apply the fixed block resampling method of Künsch (1989) so that the bootstrap replicate blocks are asymptotically iid.

4 Predictability estimates and confidence intervals

Predictability estimation using the bootstrap is based on the following two procedures:

I. Computing the baseline predictability statistic: we fit an AR-GARCH model to in-sample returns observations ($N - 22$) using standard Box-Jenkins methodology, and use the chosen model to generate conditionally optimal linear forecasts of the true data for the out-of-sample range: $1 \leq j \leq 21$, $k = 22$. These "true forecasts" from the selected model are then compared against the last 22 out-of-sample observations, and the "true", or *baseline*, value of the predictability statistic is computed. This baseline value

⁹Hsieh (1989) and Brock, Hsieh and LeBaron (1993) document nonlinear dependence in various daily exchange rates. Ding, Granger and Engle (1993), Granger and Ding (1995,1996) and Mills (1996) show that power transformations of absolute returns display much longer memory than actual returns for a variety of daily financial time series.

is calculated for each of the 22 out-of-sample horizons, and is required for the construction of bias-corrected predictability confidence intervals.¹⁰

II. The bootstrap methodology follows the 5-step procedure of Freedman and Peters (1985):

(1) As described in section 2, we use constrained maximum likelihood estimation to fit AR-GARCH models to all N observations, with the selected order of the AR-GARCH models denoted as $p(\hat{\alpha})$. The chosen models used to compute the corresponding fitted values of y_t for all $t = p(\hat{\alpha}) + 1, \dots, N$ are given by $\hat{y}_t = \mu + \sum_{i=1}^p \hat{\alpha}_i y_{t-i} + e_t$. All observations are used in order to construct pseudo-data vectors in the following steps based on $N - p(\hat{\alpha})$ errors.

(2) For each fitted model, the true in-sample error vector $\hat{e} ((N - p) \times 1)$ is the difference between the true vector \underline{y} and its conditional fitted values: $\hat{e} = \underline{y} - \hat{\underline{y}}$. The empirical distribution of \hat{e} is then resampled B times with replacement, yielding $b = 1, \dots, B$ bootstrap replicate error vectors \underline{e}^b . We use $B = 500$ replications throughout. As discussed above, applying the block resampling technique of Künsch (1989) yields asymptotically iid bootstrap replications. Each true residual vector is partitioned into 25 blocks and resampling is with replacement from the block index. The intuition here is that choosing a very large number of blocks would destroy any residual autocorrelation existing in the true data, so a moderate number is better suited.¹¹

(3) Each bootstrap error vector is used to construct a pseudo-data vector \underline{y}^b of dimension $N - p$ using the true data coefficients ($\hat{\alpha}$) for each given model. In this way we construct a total $b = 1, \dots, B$ pseudo-data vectors. The components of each are given by $y_t^b = \mu + \sum_{i=1}^p \hat{\alpha}_i y_{t-i}^b + e_t^b$, $t = p + 1, \dots, N$. Each pseudo-data vector is evaluated recursively, with the first p observations corresponding to the true returns data. These initial values should not affect the pseudo-data distribution for large sample sizes.

¹⁰Efron and Tibshirani (1993).

¹¹See Künsch (1989) and Aczel and Josephy (1992). An extension to the paper involves applying the stationary bootstrap of Politis and Romano (1994), which deals with error autocorrelation using a variable block length.

(4) From the total $N - p$ pseudo-sample returns observations, $N - k - p$ are reserved for in-sample estimation, and the last $k = 22$ are used for out-of-sample forecast evaluation. The estimation of step (1) is then repeated for each pseudo-data vector ($b = 1, \dots, B$), using the first $(N - k - p)$ observations. The AR-GARCH coefficient vector ($\underline{\alpha}^b$) is estimated as $y_t^b = \mu^b + \sum_{i=1}^p \alpha_i^b y_{t-i}^b + e_t^{*b}$, while the remaining k values are the "true" pseudo-data for each bootstrap.¹²

(5) For each pseudo-data vector, we generate $k = 22$ pseudo-forecasts, i.e. artificial predictions for the $j = 1, \dots, k$ out-of-sample horizons of interest. The forecasts are given by $y_{t+j}^{*b} = \mu^b + \sum_{i=1}^p \alpha_i^b y_{t+j-i}^{*b} + e_{t+j}^{*b}$, with $e_{t+j}^{*b} = 0$ for all j , with all future errors set to their conditional expectation of zero.

The expected MSE loss at horizon j is then just:

$$MSE_j = \frac{1}{j} \sum_{t=N-k-p+j}^{N-p} (y_{t+j}^b - y_{t+j,t}^{*b})^2$$

Relative predictability $\hat{P}(j, k)$ is obtained by computing 1 minus the ratio of expected losses at different horizons j and k for each pseudo-data vector and averaging the outcome over all B bootstrap replications:

$$\hat{P}(j, k) = 1 - \frac{E_t(L(e_{t+j,t}^b))}{E_t(L(e_{t+k,t}^b))} = 1 - \frac{MSE_j}{MSE_k} \quad (5)$$

Standard bootstrap percentile confidence intervals for the predictability statistic can be constructed directly from the percentiles of the bootstrap distribution. However, they are not reported because they are very wide for all three exchange rates. Normal or studentized confidence intervals are also not reported pending results on the asymptotic normality of \hat{P} .

Efron and Tibshirani's (1993) bias-corrected (*BC*) confidence intervals adjust the endpoints of the bootstrap percentile confidence intervals to correct for possible bias in the bootstrap statistics. This is done by comparing them to the *baseline* value of predictability, which serves as a benchmark

¹²Because of computational constraints, the lag orders p used for each pseudo-data estimation are assumed to be the same as for the true data.

for $P(j, k)$. Baseline predictability for $1 \leq j \leq 21$ and $k = 22$ is computed at stage I above by evaluating the forecasts against the true out-of-sample data using MSE. Bias correction is then measured as the proportion of bootstrap replications which are less than the baseline predictability estimate. In contrast to t -percentile confidence intervals, BC confidence intervals correct for the fact that the empirical distribution of the predictability statistic is not symmetric around baseline predictability—in this way, they have better effective coverage.¹³

5 Results and discussion

Figures 1-3 report 95% BC confidence intervals of relative predictability. The figures also show the mean of the predictability statistic for all bootstrap replications lying within the BC confidence intervals. For each fitted model, panels A and B compare the evolution of the respective exchange rate returns' intra-month predictability—evaluated under MSE—to returns predictability using a random walk. It was argued earlier that under MSE the latter is linearly declining in the forecast horizon, so it is a useful benchmark for comparing the out-of-sample performance of different models used to forecast a particular time series.

Comparing the intra-month evolution of relative predictability in Figures 1-3 suggests that $\hat{P}(\cdot)$ is higher for the £/\$, intermediate for JPY/\$ and lowest for the DM/\$ returns. All three BC confidence intervals are very wide—including negative values—and their upper (97.5%) bound exceeds predictability under the random walk. However, the mean of all bootstrap predictability statistics lying within the BC bounds lies significantly below the random walk.¹⁴ Thus, fitting AR-GARCH models we find that all 3 exchange rates are more predictable under a random walk than under the chosen model for all intra-month forecast horizons.

¹³Efron and Tibshirani's (1993) *accelerated* confidence intervals which use the jackknife were not computed as they are very computationally intensive for our dataset.

¹⁴Median predictability—not shown—lies above the mean for all three exchange rates, at about the level of the random walk. This is because the predictability statistic's distribution is positively skewed, as it is unbounded below but bounded above by 1.

Turning to discuss the evolution of the series' predictability as the forecast horizon varies, the 95% BC confidence intervals for the DM and yen/\$ actual returns are both narrower than that for the £/\$. Figure 4 compares the mean predictability, for all bootstrap replications lying within the BC bounds, of all 3 exchange rate returns' predictability. Clearly, the £/\$ exchange rate return has the higher mean predictability across intra-month forecast horizons, while the DM/\$ actual return has the lower. The £/\$'s return predictability drops sharply until 4 days-ahead, and that of the yen/\$ until 6 days-ahead. Thereafter, they decline relatively smoothly. In both cases, the "threshold" level of predictability is about 0.2.

Finally, for illustration purposes, the above methodology was also employed to estimate the relative predictability of *absolute* exchange rate returns under ARMA models and compare it to that under a random walk. Figures 1-3 show that this is much greater than that of actual returns, and the corresponding bias-corrected confidence intervals narrower. This may be due to the significantly lower explanatory power of parametric models for modeling actual as opposed to absolute returns. The smaller parameter uncertainty for absolute returns thus results in a narrower range for the corresponding bootstrap predictability estimates. Absolute returns of the yen/\$ exchange rate are found to have relatively higher predictability, suggesting that the volatility of yen/\$ returns has been more stable compared to other exchange rates.

6 Conclusion

This paper developed the Diebold-Kilian (1997) measure of relative time series predictability and applied it to daily exchange rate returns for 3 currencies. Because of the nonnormality of the predictability statistic, bootstrap inference methods were used. AR-GARCH models were estimated for exchange rate returns, and forecasts evaluated according to MSE loss. It was found that exchange rate returns are less predictable using the model than using a random walk beyond 1-week forecast horizons for all 3 currencies. Overall, the results confirm the greater difficulty that parametric models have in forecasting daily actual exchange rate returns. Against that background, the predictability statistic allows for a concrete evaluation of the difference

in alternative models' expected losses at different forecast horizons.

The robustness of our results could be assessed by varying the in-sample and out-of-sample periods, as well as by changing the size of the full dataset itself. Researchers have often used very long data sets for model-fitting without regard to their out-of-sample performance. However, the latter would likely indicate a more selective approach, such as applying lower weights in parameter estimation to distant data. Examples of models with very long daily samples include Ding, Granger and Engle (1993), Hentschel (1995) and Pesaran and Timmerman (1995). An exception is Satchell and Timmerman (1995), who use a recursive sample of only the 1000 most recent data. We speculate that an information-theoretic selection criterion may be available. These extensions are the subject of current research.

References

- [1] Aczel, A. and N. Josephy 1992, "Using the Bootstrap for Improved ARIMA Model Identification", *Journal of Forecasting* 11, 71-80.
- [2] Bera, A. and Higgins, M. 1995, "On ARCH Models: Properties, Estimation and Testing", in Oxley, L., George, D., Roberts, C. and S. Sayer (eds), *Surveys in Econometrics*, Oxford: Blackwell.
- [3] Beran, J. 1994, *Statistics for Long-Memory Processes*. New York: Chapman and Hall.
- [4] Berkowitz, J. and L. Giorgianni 1997, "Long-Horizon Exchange Rate Predictability?", *IMF Working Paper* WP/97/6.
- [5] Brock, A., Hsieh, D. and B. Lebaron 1993, *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*. Cambridge, MA: The MIT Press.
- [6] Brooks, C. 1997, "Linear and Non-linear (Non-)Forecastability of High-frequency Exchange Rates", *Journal of Forecasting* 16, 125-145.
- [7] Christoffersen, P. and F. Diebold 1996, "Further Results on Forecasting and Model Selection Under Asymmetric Loss", *Journal of Applied Econometrics* 11, 561-572.
- [8] Christoffersen, P. and F. Diebold 1997, "Optimal Prediction Under Asymmetric Loss", NBER Technical Working Paper No.167, forthcoming in *Econometric Theory*.
- [9] Diebold, F. 1998, "The Past, Present and Future of Macroeconomic Forecasting", *Journal of Economic Perspectives* 12, 175-192.
- [10] Diebold, F. and L. Kilian 1997, "Measuring Predictability: Theory and Macroeconomic Applications", NBER Technical Working Paper 213.
- [11] Diebold, F. and R. Mariano 1995, "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics* 13, 253-263.

- [12] Diebold, F. and J. Nason 1990, "Nonparametric Exchange Rate Prediction?", *Journal of International Economics* 28, 315-332.
- [13] Ding, Z., Granger, C. and R. Engle 1993, "A Long-Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance* 1, 83-106.
- [14] Efron, B. and R. Tibshirani 1986, "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy", *Statistical Science* 1, 54-77.
- [15] Efron, B. and R. Tibshirani 1993, *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- [16] Freedman, D. and S. Peters 1985, "Using the Bootstrap to Evaluate Forecasting Equations", *Journal of Forecasting* 4, 251-262.
- [17] Granger, C. and Z. Ding 1995, "Some Properties of Absolute Return: An Alternative Measure of Risk", *Annales d'Economie et de Statistique* 40, 67-92.
- [18] Granger, C. and Z. Ding 1996, "Varieties of Long Memory Models", *Journal of Econometrics* 73, 61-78.
- [19] Granger, C. and P. Newbold 1986, *Forecasting Economic Time Series*. Orlando: Academic Press.
- [20] Hall, P. 1992, *The Bootstrap and Edgeworth Expansion*. New York: Springer Verlag.
- [21] Hall, P. and J. Horowitz 1996, "Bootstrap Critical Values for Tests Based on Generalized Methods of Moments Estimators", *Econometrica* 64, 891-916.
- [22] Hentschel, L. 1995, "All in the Family: Nesting Symmetric and Asymmetric GARCH Models", *Journal of Financial Economics* 39, 71-104.
- [23] Hsieh, D. 1989, "Modeling Heteroscedasticity in Daily Foreign-Exchange Rates", *Journal of Business and Economic Statistics* 7, 307-317.
- [24] Künsch, H. 1989, "The jackknife and the bootstrap for general stationary observations", *Annals of Statistics* 17, 1217, 1241.

- [25] Masarotto, G. 1990, "Bootstrap Prediction Intervals for Autoregressions", *International Journal of Forecasting* 6, 229-239.
- [26] Meese, R. and K. Rogoff 1983, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?", *Journal of International Economics* 14, 3-24.
- [27] Mills, T. 1996, "Non-linear Forecasting of Financial Time Series: An Overview and Some New Models", *Journal of Forecasting* 15, 127-135.
- [28] Pesaran, R. and A. Timmerman 1995, "Predictability of Stock Returns: Robustness and Economic Significance", *Journal of Finance* 50, 1201-28.
- [29] Politis, D. and J. Romano 1994, "The Stationary Bootstrap", *Journal of the American Statistical Association* 89, 1303-1313.
- [30] Rothemberg, J. and M. Woodford 1996, "Real Business-Cycle Models and the Forecastable Movements in Output, Hours and Consumption", *American Economic Review* 86, 71-89.
- [31] Satchell, S. and A. Timmerman 1995, "An Assessment of the Economic Value of Nonlinear Foreign Exchange Rate Forecasts", *Journal of Forecasting* 14, 477-97.
- [32] Souza, R. and A. Neto 1996, "A Bootstrap Simulation Study in ARMA(p, q) Structures", *Journal of Forecasting* 15, 343-353.

Table 1: GBP/US\$ Returns**A. Summary statistics¹⁵**

Mean	0.0000446
Standard Deviation	0.0062
Skewness	0.02780
Kurtosis	5.6367
Normality	810.24 (0.00)
ADF(5)	-20.76 ^c
$Q_x(20)$	46.23 (0.001 ^c)

B. Model specification¹⁶

AR	7
MA	0
GARCH(p,q)	(1,1) / GARCH-M
ARMA errors	No
SIC	-7.4408
$\overline{R^2}$	0.0119
$Q_x(20)$	18.90 (0.53)
ARCH LM (20)	1.22 (0.22)

¹⁵ $N = 2,678$ observations. ADF(5) is a unit root test with 5 lags. Normality is the Bera-Jarque test, asymptotically distributed $\chi^2(2)$. $Q_x(20)$ is the Ljung-Box statistic of order 20. ARCH(5) is Engle's LM test for ARCH.

¹⁶The order p of the AR components was obtained using AIC and SIC. $\overline{R^2}$ is adjusted R^2 . ARCH reports the F-statistics and P-values for the ARCH LM test of no linear dependence in the squared errors.

Table 2: DM/US\$ Returns**A. Summary statistics¹⁷**

Mean	0.0000577
Standard Deviation	0.0066
Skewness	0.0328
Kurtosis	5.0023
Normality	447.83
ADF(5)	-21.84 ^c
$Q_x(20)$	22.21 (0.33)

B. Model specification¹⁸

AR	11
MA	0
GARCH(p,q)	(1,1)
ARMA errors	No
SIC	-7.2623
$\overline{R^2}$	0.004572
$Q_x(20)$	22.33 (0.32)
ARCH LM (20)	1.53 (0.06 ^a)

¹⁷ $N = 2678$ observations. ADF(5) is a unit root test with 5 lags. Normality is the Bera-Jarque test, asymptotically distributed $\chi^2(2)$. $Q_x(80)$ is the Ljung-Box statistic of order 20. ARCH(5) is Engle's LM test for ARCH.

¹⁸The order of the AR components was obtained using AIC and SIC. $\overline{R^2}$ is adjusted R^2 . ARCH reports the F-statistics and P-values for the ARCH LM test of no linear dependence in the squared errors.

Table 3: JPY/US\$ Returns**A. Summary statistics¹⁹**

Mean	0.000036
Standard Deviation	0.0066
Skewness	-0.2346
Kurtosis	6.1223
Normality	1112.36
ADF(5)	-21.68 ^c
$Q_x(20)$	26.05 (0.16)

B. Model specification²⁰

AR	5
MA	0
GARCH(p,q)	(1,1)
ARMA errors	(1,1)
SIC	-7.264
$\overline{R^2}$	0.0077
$Q_x(20)$	15.34 (0.64)
ARCH LM (20)	0.39 (0.99)

¹⁹ $N = 2,678$ observations. ADF(5) is a unit root test with 5 lags. Normality is the Bera-Jarque test, asymptotically distributed $\chi^2(2)$. $Q_x(80)$ is the Ljung-Box statistic of order 20. ARCH(5) is Engle's LM test for ARCH.

²⁰The order of the AR components was obtained using AIC and SIC. $\overline{R^2}$ is adjusted R^2 . ARCH reports the F-statistics and P-values for the ARCH LM test of no linear dependence in the squared errors.