Predicting Spot Exchange Rates in a Nonlinear Estimation Framework Using Futures Prices*

A. M. PARHIZGARI
MARIA EUGENIA DE BOYRIE

INTRODUCTION

The past two decades have seen enormous growth in the literature on foreign exchange rates. Following the adoption of floating exchange rates by the major industrial trading nations in the early 1970s, a wealth of data on exchange rates and on the factors that determine them have been generated. This amount of data has enabled researchers to test a variety of hypotheses and theories on the foreign exchange markets.

Despite considerable research, many unresolved issues remain. The major theories on exchange-rate determination center on the asset market, the monetary market, and the balance of payments as determinants, although models based on these theories have not been fully successful in explaining exchange rate movements. To some extent, the problem is

*An earlier version of this article was presented at the October 1995 Financial Management Association (FMA) meetings in New York. The authors are grateful to Professor Claire G. Gilmore of St. Joseph's University and other participants at the meetings for their helpful comments and suggestions. They are also greatly indebted to the journal's editor and two anonymous referees for their constructive comments. All remaining errors, of course, are the authors' alone.

- *A. M. Parhizgari is a Professor of Finance and International Business at Florida International University.*
- *Maria Eugenia de Boyrie is an Assistant Professor of Finance at Florida International University.*

The Journal of Futures Markets, Vol. 17, No. 8, 935–956 (1997)

CCC 0270-7314/97/080935-22

that exchange rate movements depend on expectations of the future path of the exchange rate, which, in turn, depends on expectations of relative inflation and growth rates, among other fundamentals. Such expectations are highly volatile and short-lived, making them difficult to model. It is thus not surprising that most of the developments in the exchange-rate literature over the last decade or so have been concerned with the use of the appropriate techniques to capture the expectation variables.

The need for accurate exchange-rate forecasting is more pressing today with the constant volatility of currency values and a growing volume of international transactions. At the same time, there is no consensus as to which model or type of model can provide a more accurate exchange-rate forecast.

The purpose of this study, therefore, is to construct and test several forecasting models with the aim of improving upon and extending the prior results. The models range from simple to complex specifications and employ a variety of estimation techniques, including ordinary least-squares regression (OLS), locally weighted regression, and neural networks.

Section II presents the literature and background research. The structure of the proposed forecasting model is developed in Section III. Section IV describes the sources of data, explains some refinements to the data, and presents the empirical results. Section V contains a summary and conclusion.

LITERATURE REVIEW AND BACKGROUND RESEARCH

A vast amount of literature tests the in-sample forecasting ability of both the monetary approach and the portfolio-balance approach models. In the early 1980s, work in this area seemed futile: Haynes and Stone (1981), Frankel (1984), and Backus (1984) concluded that estimates of the real interest rate differential could not explain in-sample exchange-rate changes.

Rasulo and Wilford (1980) and Haynes and Stone (1981) attribute the poor performance of the monetary approach to the constraints imposed on relative monies, income, and interest rates. Frankel (1982) provides an alternative explanation for the poor performance of the monetary models. He introduces home and foreign wealth into the money demand equations, and ignores any constraints on domestic and foreign income, wealth, and inflation terms. His results show that his monetary approach equation fits the data well and estimates the signs correctly.

Similar attempts to explain the poor performance of the monetary models and improve on the results obtained in previous studies are by Driskill and Sheffrin (1981), Boughton (1988), Hoffman and Schlagenhauf (1983), and McDonald and Taylor (1991). Their results have not been completely convincing.

Meese and Rogoff (1983), in their seminal study, attempt to determine how well exchange-rate models perform in estimating out-of-sample forecasts. They test the flexible-price model, the real interest differential model, and the portfolio monetary synthesis of Hooper and Morton (1982). They conclude that none of the exchange-rate models using the asset approach can really outperform the simple random-walk model.

In a follow-up study, Meese and Rogoff (1984) impose coefficient constraints and reestimate their models. Once again, their model is not able to outperform the random walk in the short run.

Since the work of Meese and Rogoff (1983, 1984), most of the literature has been devoted to determining whether their specification of the asset reduced-form model, their estimation strategy, or the models themselves are at fault. Boughton (1987), Woo (1985), Finn (1986), Somanath (1986), and Schinasi and Swamy (1989) have replicated the work of Meese and Rogoff (1984) by improving upon and estimating different versions of the rational-expectations form of the flexible-price model that include a partial adjustment term in money demand. Their work contributes to the literature in that they all successfully identify models that at times perform better than a random walk.

Whereas the aforementioned studies have primarily employed linear specifications, research in the late 80s and early 90s has concentrated on nonlinear foreign exchange models [see, for example, Hsieh (1989), Diebold and Nason (1990), Engel and Hamilton (1990), Chinn (1991), Kuan and Liu (1992), McGuirk, Robertson, and Spanos (1993), and Engel (1994)].

Hsieh (1989) tests for nonlinear dependence in foreign exchange rates by applying the Brock, Dechert, and Scheinkman (BDS) test. This test detects strong nonlinear dependence. In particular, large and small changes in returns tend to be systematically clustered together when plotted over time. Friedman and Vanderstell (1982), Diebold and Nerlove (1989), Bollerslev (1987), and Baillie and Bollerslev (1989) confirm the nonlinear temporal dependence of exchange-rate returns.

Some of the most recent empirical and theoretical results also support nonlinearity in the exchange rates [see Engel and Hamilton (1990), Chinn (1991), and McGuirk et al. (1993)].[1] Diebold and Nason (1990),

[1]Other authors who have researched the nonlinearity evidence are Diebold (1988), Diebold and Nerlove (1989), Diebold and Nason (1990), Hsieh (1988), and Meese and Rose (1991).

however, conclude that nonlinearities of exchange rates cannot be exploited to generate improved point predictions relative to linear models.

Forecasting nonlinear structural models is complicated and normally cannot be accomplished without a base model. There have been several attempts in this area. For example, Diebold and Nason (1990) use a weighted nearest neighbor technique [also known as locally weighted regression (LWR)] to forecast exchange rate changes. Engel and Hamilton (1990) employ a segmented trend model, wherein a nonstationary time series is decomposed into a sequence of stochastic segmented time trends. Chinn (1991) and Chinn and Frankel (1994) use an alternating conditional expectations (ACE) technique wherein they transform all the variables to obtain a linearized relationship between them. In general, these efforts demonstrate that nonlinear specifications do somewhat better than most linear models at the four-quarter horizon.

Finally, the use of artificial-intelligence techniques in modeling exchange rates began fairly recently. For example, the Kuan and Liu (1992) alternative to forecasting exchange rates uses both feed-forward and recurrent neural networks. They compare their results with those obtained with the use of ARMA models for in-sample, out-of-sample, and sign predictions. As in Diebold and Nason (1990), they conclude that nonlinearities of exchange rates may not be exploited to improve point prediction. Looking exclusively at the sign predictions, their results show that the neural-network models perform better than the ARMA models.

To sum up, there is no clear picture when it comes to using nonlinear time series techniques. The evidence indicates that nonlinear models can outperform a random walk in sample, but not forecasting out of sample.

METHODOLOGY

The above brief examination of the literature on forecasting foreign exchange demonstrates that neither theoretically structured asset market models nor ad hoc nonlinear models offer consistent improvement over the forecasting performance of the random walk model.

Most of the asset market models that are tested include variables such as interest rates, home and foreign wealth, inflation, cumulated trade balance, real income, money supply, balance of payments, and gross national product (GNP). Most of these variables have been shown to have little significance in explaining changes in exchange rates. In ad hoc nonlinear models, one major variable is used: the spot rate. These models seem to fare better than the asset-market models. In some cases, the nonlinear models outperform the random-walk model for in-sample predictions; this is not the case for out-of-sample forecasts.

These results reveal a need for a model, or a set of models, that can better explain the changes in foreign exchange rates. This study uses the futures market, an established financial arena, which, by definition, projects beyond the current time period (= ex ante) in its entirety. This allows the development of a model that includes variables previously not used. One such variable is the price of futures contracts on foreign exchange.

It is postulated that the wealth of information embedded in futures contracts on foreign exchange rates reflects the market's expectation of expected future exchange rates. Thus, future spot rates of a currency are hypothesized to depend upon the market's current expectation of the price of a set of futures contracts on that currency in the futures currency market; that is

$$s_0^t = f\,(P_0^{t'}, P_0^{t''}, P_0^{t'''}, \ldots) \tag{1}$$

where

s_0^t = the current (= subscript 0 in notations) expected spot rate of a currency t periods ahead

and

$P_0^{t's}$ = the current market price of futures contracts on the same currency with maturity at t', t'', t''', et cetera ($t < t' < t'' < t'''$).

Under the *market efficiency hypothesis* relationship (1) may be further simplified to include only the prices of more recent futures contracts with the maturity close to t, which is the time period for the spot rate to be predicted. Thus, $P_0^{t'}$ might be a sufficient statistic for the information conveyed by the futures market in the sense that it contains the most recent and relevant information regarding expected future spot rates.

Literature on the *market efficiency hypothesis* abounds. The hypothesis, though undisputed at the *theoretical* front, has raised some questions at the *empirical* level, leaving room for the incorporation of further futures prices. Though there is an increasing body of work that tends to support market efficiency, the problem does not seem to be fully resolved [see, for example, Chowdhury (1991, p. 587), Schwartz and Laatsch (1991, pp. 669–670 and 682)]. With the use of more recent methods such as unit roots and co-integration, a number of studies are still investigating the degree and sources of inefficiency. Some ambiguities still remain. For example, although co-integration of the spot price and futures price is used to demonstrate market efficiency [see Hakkio and Rush

(1989) and Shen and Wang (1990)], Granger (1986) shows that co-integration between prices in two different markets implies inefficiency.[2]

Note that relationship (1) is stated with a functional operator. The underlying design of this relationship, which is discussed in some detail in Appendix A, is to provide forecasts beyond the current time period (=ex ante) at any time. Further specification of this relationship will depend upon the methodology employed at the estimation level. Additive linear and nonlinear forms as well as free forms are among the alternatives that are considered.

Relationship (1), which attempts to forecast spot rates solely on the basis of *futures prices* (FP), is henceforth referred to as the FP model. One advantage of this model is its relative simplicity. It uses data that are available, especially for out-of-sample prediction. Such data for purely ex ante forecast are either unavailable or difficult to obtain in models based on the fundamentals.

If one draws upon prior empirical results (see Section II), consideration of the past history of spot rates in forecasting is very compelling. As an alternative extension, relationship (1) is augmented to include past spot rates. The resultant model, which thus attempts to forecast spot rates on the basis of *futures prices* (FP) and *spot prices* (SP), is henceforth referred to as FPSP.

A direct comparison of the results obtained from the above two models is made with the random-walk (hereafter RW) model to gauge their performance. The random-walk model assumes that the spot prices at time t are the best unbiased estimators of the future spot. The RW model thus forecasts exchange rates on the basis of only past *spot prices* (SP).

The above three models are estimated and then forecasted under variant techniques: ordinary least squares (OLS), locally weighted regressions (LWR), and neural networks (NN). Details of the first technique can be found in most standard texts in econometrics. A preliminary review of the other two techniques is presented in Appendix B. Further details of these latter techniques may be found in Cleveland, Devlin, and Grosse (1988) and Diebold and Nason (1990) for LWR, and in Caudill (1989) and Kuan and Liu (1992) for neural networks.

The accuracy of both in-sample and out-of-sample forecasts is measured by two statistics: mean-squared error (MSE), and root-mean-squared percentage error (RMSPE).

[2]Pertinent to this latter point, see tests by Kawaller and Koch (1984), Yadav and Pope (1991), and MacDonald and Hein (1993).

DATA AND EMPIRICAL APPLICATION

Data on the futures exchange prices are obtained from the Futures Industry Institute Data Center. The sample period chosen is from January 2, 1985 through June 30, 1994. The futures prices data used in this analysis consist of daily settlement prices on contracts for the British pound (BP), Canadian dollar (CD), German mark (DM), Japanese yen (JY), and the Swiss franc (SF). These have been the most active currencies in daily futures trading during the 1985–1994 period. Each currency is traded in four contracts: March, June, September, and December.

Both spot and the futures prices are converted into rates of change. The rate of change is calculated by taking the difference of the logarithms of two consecutive daily rates, that is, $[\ln(P_{t+1}) - \ln(P_t)]$, where P_t is the spot or the futures price contract at time t. An advantage of using logarithmic changes is that it removes (or mitigates) problems associated with the estimation of nonstationary nonparametric regression functions, as well as numerical problems with highly collinear regressors.[3]

The data are scanned for each currency to locate days on which three or more contracts are traded. For forecasting purposes, the sample period is broken into two subperiods. The first is the estimation period, which uses the sample period from January 2, 1985 through December 30, 1993, and the second is the forecasting period, which uses the sample period from January 2, 1994 through June 30, 1994. After disregarding the days for which a spot price is not recorded, the number of observations per currency for the in-sample estimation is reduced to: (1) British pound, 1,989 observations; (2) Canadian dollar, 2,014; (3) German mark, 2,046; (4) Japanese yen, 2,078; and (5) Swiss franc, 2,073. The number of observations used in the out-of-sample forecast is: (1) British pound, 249; (2) Canadian dollar, 248; (3) German mark, 250; (4) Japanese yen, 251; and (5) Swiss franc, 249.

Empirical application of the *three* models (FP, FPSP, and RW, see Section III) for *five* currencies, under *three* distinct techniques (OLS, LWR, and NN), and across *six* designated subsample periods[4] generates massive output, too large for inclusion in this article. The preliminary

[3]Studies by Mussa (1979) and Doukas and Rahman (1987) show that the log changes of future prices follow a stationary distribution. That is, the process generating the log of currency futures rates is well approximated by a random walk.

[4]To test for stationarity, the in-sample period is further divided into six subsamples for estimation. The first subsample covers approximately 10 years of data, from January 2, 1985, through December 30, 1993. The remaining five subsamples are obtained by rolling a sample period of 4 years forward. With the use of each of these five subsamples, a series of Chow tests to examine the stability of the coefficients are performed. In general, the results show that parameter stability is not rejected at the 5% level.

runs, however, reveal that the 1990–1993 subperiod offers the best performance relative to the other earlier subperiods in both in-sample and out-of-sample prediction, and that the choice of model is *currency, time,* and *criteria* dependent. Furthermore, the models estimated on the 1990–1993 subperiod provide better forecast accuracy than when the entire 10 years of data are used. As may be expected, a more recent data group is more representative because it includes more up-to-date market information. This is true for all the currencies studied. To this end and to preserve space, only summary results are reported on the outcome of the 1990–1993 sample period for estimation and the 1993–1994 sample period for out-of-sample forecasting.[5] Some brief technical notes pertaining to a few other facets of the empirical applications are included in Appendix A.

For in-sample prediction the model that minimizes the MSE consistently is the FPSP model, that is, the model that uses *both* the past spot rates and the futures prices to forecast exchange rates. When RMSPE is used as the selection criterion, the RW model seems to perform the best.

For *out-of-sample* prediction the results are mixed. A summary of the outcomes across model–currency criteria and forecasting horizon for the 1990–1993 subsample period is provided in Table I and an overall summary is given in Table II. Further details upon which these two tables are based are included in Tables III–V.[6] To facilitate the link between Tables I and II and Tables III–V, in each of the latter tables, MSE or RMSPE that designates the best choice of the model for *each currency* is boxed in rectangles.

Table I summarizes the results along three dimensions: currency, technique, and forecasting horizon. The fourth dimension of the study, that is, selection criterion, is portrayed in Panels A and B. The models reported in this table offer the best performance across the above four dimensions. This table shows that accuracy in forecast depends not only on the forecast horizon, but also on the choice of model, currency, and to some extent the selection criterion. For example, based on the *first* row of Table I, the best model that forecasts the German mark under the OLS technique is the FPSP model for the 3-month horizon. Similarly, the best model that forecasts the Japanese yen is the FPSP model, but for the 1-month horizon. If one continues with the case of DM with MSE as the criterion, under the LWR and NN techniques, the RW model performs

[5]Tabulated details of other subperiods are available from the authors.

[6]These tables may be viewed as representative of the extensive searches that have been performed. The footnotes to these tables provide sufficient explanatory information to link each table with the reported outcomes in Tables I and II. The MSEs (RMSPEs) that yield the selected models are *boxed*. The outcomes reported in Tables I and II are based on these selections.

TABLE I

Selected Models Based on the Best Out-of-Sample Performance for Each Currency under Different Techniques and Horizons

Currency Technique	*BP Horizon (in Months)*			*CD Horizon (in Months)*			*DM Horizon (in Months)*			*JY Horizon (in Months)*			*SF Horizon (in Months)*		
	1	*3*	*6*	*1*	*3*	*6*	*1*	*3*	*6*	*1*	*3*	*6*	*1*	*3*	*6*
Panel A: Selection Criterion–Mean-Squared Errors (MSE)															
OLS	–	RW	–	–	RW	–	–	FPSP	–	FPSP	–	–	–	–	FP
LWR	FP	–	–	RW	–	–	–	–	RW	FP	–	–	–	–	RW
NN	RW	–	–	–	RW	–	–	RW	–	RW	–	–	FPSP	–	–
Panel B: Selection Criterion—Root-Mean-Squared Percentage Errors (RMSPE)															
OLS	–	–	RW	–	–	FP	–	FP	–	–	RW	–	–	FPSP	–
LWR	FPSP	–	–	FP	–	–	–	–	RW	FP	–	–	–	FPSP	–
NN	RW	–	–	–	RW	–	–	RW	–	FPSP	–	–	FP	–	–

Notes: FP = the model that uses *futures prices* only [see eq. (A.1)], PFSP = the model that uses *futures prices* and *spot prices*, and RW = random-walk model. BP = British pound, CD = Canadian dollar, DM = German mark, JY = Japanese yen, SF = Swiss franc. OLS = ordinary least-squares regression, LWR = locally weighted regression, NN = neural networks. The *boxed* items under DM are used as examples in the text. They are also used to portray links between this table and Table II.

TABLE II

Summary Results—Final Model Selection Based on the Best Out-of-Sample Performance across Different Techniques

Currency	*BP*	*CD*	*DM*	*JY*	*SF*
Panel A: Selection Criterion—Mean-Squared Errors (MSE)					
Model	RW	RW	FPSP	FPSP	FPSP
Technique	NN	LWR	OLS	OLS	NN
Panel B: Selection Criterion—Root-Mean-Squared Percentage Errors (RMSPE)					
Model	FPSP	FP	RW	FP	FPSP
Technique	LWR	LWR	NN	LWR	OLS

Note: FP = the model that uses *futures prices* only [see eq. A.1, Appendix A], FPSP = the model that uses *futures prices* and *spot prices*, and RW = random-walk model. BP = British pound, CD = Canadian dollar, DM = German mark, JY = Japanese yen, SF = Swiss franc. OLS = ordinary least-square regression, LWR = locally weighted regression, NN = neural networks. The *boxed* items under DM are used as examples in the text. They are also used to portray links between Tables I and II.

TABLE III

Summary Statistics—Out-of-Sample MSE and RMSPE under the OLS Technique

		Mean-Squared Errors (MSE)			*Root-Mean-Squared Percentage Errors (RMSE)*		
Model	*Currency*	*Forecasting Horizon 1 Month Ahead*	*Forecasting Horizon 3 Months Ahead*	*Forecasting Horizon 6 Months Ahead*	*Forecasting Horizon 1 Month Ahead*	*Forecasting Horizon 3 Months Ahead*	*Forecasting Horizon 6 Months Ahead*
RW	BP	0.11203	$\boxed{0.10640}$	0.17185	1.19547	1.26020	$\boxed{1.03029}$
	CD	0.09059	$\boxed{0.06445}$	0.08145	0.92385	1.07285	1.10204
	DM	0.40633	0.27280	0.33070	1.01366	1.10068	1.04690
	JY	0.22141	0.40988	0.38832	1.04283	$\boxed{1.02784}$	1.14077
	SF	0.27980	0.44060	0.32686	1.02347	0.95943	1.32912
FP	BP	0.14299	0.11081	0.18126	1.26275	1.43756	1.11344
	CD	0.09143	0.09059	0.08277	0.90378	1.82275	$\boxed{1.04866}$
	DM	0.40770	0.26967	0.31603	1.06400	$\boxed{0.97891}$	1.06311
	JY	0.23104	0.42291	0.52032	1.07491	1.07702	1.58348
	SF	0.71027	0.52116	$\boxed{0.27552}$	1.00175	0.96385	1.03268
FPSP	BP	0.14426	0.11343	0.18122	1.24592	1.57637	1.15258
	CD	0.09640	0.11509	0.07877	1.06057	2.44902	1.06325
	DM	0.40834	$\boxed{0.26777}$	0.31844	1.09362	1.03053	1.03899
	JY	$\boxed{0.21895}$	0.43408	0.50642	1.08208	1.18417	1.39553
	SF	0.27619	0.43297	0.35421	1.05728	$\boxed{0.95439}$	1.28000

Note: MSE values are $\times 10^4$ and RMSPE's are $\times 10^2$. FP = the model that uses *futures prices* only, FPSP = the model that uses *futures prices* and *spot prices*, and RW = random walk model. BP = British pound, CD = Canadian dollar, DM = German mark, JY = Japanese yen, SF = Swiss franc. The MSEs (RMSPEs) that are *boxed* yield the selected models that are indicated in Table I for the *OLS technique only*. For example, under MSE for the BP currency (which is in the first row of each of the models) the minimum MSE among the nine figures is 0.10640, which corresponds with the RW model under the 3-month forecasting horizon. Thus, the RW model under the 3-month forecasting horizon is selected as the best model and is reported in Table I, Panel A (=MSE), for the BP currency (under the 3-month forecasting horizon), and the *technique used*, that is, *OLS*, is reflected.

TABLE IV

Summary Statistics—Best Out-of-Sample MSE and RMSPE under the LWR Technique

Model	*Currency*	*Smoothing Parameter* (ξ)	Forecasting Horizon in Months	Mean-Squared Error (MSE)	Root-Mean-Squared Percentage Error (RMSPE)
RW	BP	0.4	1	0.13049	3.17622
	CD	0.9	1	$\boxed{0.03988}$	1.32240
	DM	0.9	6	$\boxed{0.27714}$	$\boxed{1.13334}$
	JY	0.9	6	0.29614	0.92015
	SF	0.9	6	$\boxed{0.48517}$	0.98725
FP	BP	0.9	1	$\boxed{0.10587}$	1.80385
	CD	0.9	1	0.04122	$\boxed{0.87797}$
	DM	0.8	1	0.32159	1.25486
	JY	0.8	1	$\boxed{0.29416}$	$\boxed{0.089754}$
	SF	0.8	1	0.49221	0.99547
FPSP	BP	0.9	1	0.27982	$\boxed{0.89179}$
	CD	0.9	1	0.07898	2.28688
	DM	0.9	1	0.30480	1.96541
	JY	0.9	6	0.30146	1.20548
	SF	0.8	3	0.51042	$\boxed{0.11104}$

Note: The table reports the *best* outcome of the smoothing parameter (ξ) and the forecasting horizon. MSE values are $\times 10^4$ and RMSPE's are $\times 10^2$. FP = the model that uses *futures prices* only, FPSP = the model that uses *futures prices* and *spot prices,* and RW = random walk model. BP = British pound CD = Canadian dollar, DM = German mark, JY = Japanese yen, SF = Swiss franc. The MSEs (RMSPEs) that are *boxed* yield the selected models that are indicated in Table I for the *LWR technique* only. For example, under MSE for the BP currency (which is in the first row of each model) the minimum MSE among the three figures is 0.10587, which corresponds with and selects the FP model under the 1-month forecasting horizon as the best model. This is reported in Table I, Panel A (=MSE), for the BP currency (under 1-month forecasting horizon), in the row designating the *LWR technique.*

the best, but for different forecasting horizons (6 months and 3 months, respectively).

Disregarding some exceptions, Table II reveals a set of overall general conclusions that could be drawn on the basis of all the experiments across model–technique–currency. This table, which is extracted from Table I, further summarizes the results in terms of the best model across the *three* techniques. Considering each currency, the best model across the *three* techniques is obtained by minimum MSE (Panel A), or minimum RMSPE (Panel B). For example, for DM in Panel A of Table I—DM is boxed for ease of presentation—a comparison of the MSEs of the three selected models, that is, FPSP, RW, and RW, yields the minimum for the FPSP model. Considering that in this case the corresponding technique

TABLE V

Summary Statistics—Best Out-of-Sample MSE and RMSPE under the Neural Network Technique

Model	*Currency*	*Forecasting Horizon in Months*	*Mean-Squared Errors (MSE)*	*Root-Mean-Squared Percentage Errors (RMSPE)*
RW	BP	1	$\boxed{0.09000}$	$\boxed{0.98339}$
	CD	3	$\boxed{0.06000}$	$\boxed{0.97588}$
	DM	3	$\boxed{0.28000}$	$\boxed{0.92609}$
	JY	1	$\boxed{0.22000}$	0.96555
	SF	1	0.44000	1.03067
FP	BP	3	0.10000	1.35129
	CD	3	0.09000	1.00052
	DM	3	0.29000	0.96555
	JY	1	0.42000	0.97397
	SF	1	0.43000	$\boxed{0.99231}$
FPSP	BP	1	0.10000	1.35309
	CD	3	0.06000	0.98865
	DM	3	0.32000	1.06502
	JY	1	0.42000	$\boxed{0.94200}$
	SF	1	$\boxed{0.25000}$	1.00426

Note: The table reports the *best* outcome for the forecasting horizon. See also Appendix A, section d. MSE values are $\times 10^4$ and RMSPE's are $\times 10^2$. FP = the model that uses *futures prices* only, FPSP = the model that uses *futures prices* and *spot prices*, and RW = random-walk model. BP = British pound, CD = Canadian dollar, DM = German mark, JY = Japanese yen, SF = Swiss franc. The MSEs (RMSPEs) that are *boxed* yield the selected models that are indicated in Table I for the *NN technique* only. For example, under MSE for the BP currency (which is in the first row of each model) the minimum MSE among the three figures is 0.09000, which corresponds with and selects the RW model on the 1-month forecasting horizon as the best model. This is reported in Table I, panel A (=MSE), for the BP currency (under 1-month forecasting horizon), in the row designating the *NN technique*.

is OLS, these two items, that is, FPSP and OLS, are portrayed under DM in Panel A of Table II. Similarly, this model, that is, the FPSP model, outperforms all the other choices for the Japanese yen and for the Swiss franc as well (see Panel A). It should also be obvious from Table II that a decision based on RMSPE may yield a slightly different set of choices. For example, the deutsche mark is forecast best with the RW model under this criterion.

Under MSE and RMSPE combined, the FP and FPSP models seem to fare better than the RW model. In sum, these two models offer the best performance in nearly 70% of the cases (a total of 7 out of 10 cases in Table II).

SUMMARY AND CONCLUSION

Forecasting foreign exchange has fascinated the financial world for several decades. Finding a model that best forecasts changes in foreign exchange rates has been a challenge. This study extends the research into new avenues. It examines three models. The first model (FP) attempts to forecast future spot rates by using only prices of futures contracts on foreign exchange. The second model (FPSP) uses the prices on futures contracts and the past history of spot prices. The third is the well-known random-walk model (RW). Three different techniques are applied to each model: ordinary least squares (OLS), locally weighted regression (LWR), and neural networks (NN).

The amount of statistical information generated is considerable. For this reason, the final conclusions of this study are based mainly on results obtained with the use of the data for the 1990–1993 (estimation) and 1993–1994 (forecasting) subperiods. These subperiods, based on a set of tests on parameter stability, are as good a representation of the data as any other subperiods analyzed. In addition, these subperiods contain the most recent information available in the market.

Out-of-sample forecasts reveal that the choice of model and technique is, if not complex, an elaborate process indeed. There seems to be no best model to forecast over *all* horizons for *all* currencies on the basis of *all* selection criteria. If one compares the results across the models and techniques, and considers *all* the subperiods and forecasting horizons, a comfortable generalization, however, may be drawn. Overall, within this comprehensive setting, it could be concluded that the best model for out-of-sample forecasting appears to be the FPSP, and the best technique seems to be ordinary least squares. When this technique is used, the best forecasting horizon is 30 days ahead. This is as expected, given that in most cases the accuracy of forecasts declines as the forecasting horizon is extended. In terms of relative performance, the FPSP model under OLS is followed by neural networks and then by the locally weighted regression technique. This is evident from MSE and RMSPE criteria. There are, of course, some ad hoc exceptions to this general conclusion. For example, even though both the neural networks and the locally weighted regression techniques perform better out of sample than in sample, they do not always outperform OLS.

A few anomalies are observed. First, in some cases, the out-of-sample forecasting performance of the models is better than their in-sample results. Second, in these cases and/or in some other cases, forecasts for the longer-term horizons are better than for the shorter terms. These two points, which are mostly attributes of the FP and FPSP models, are somewhat curious, because forecasting models, in general, perform better in sample than out of sample, and for shorter horizons than for longer forecasts. It may be that this can be attributed to the information content of the futures prices data in the sense that the currency futures market during the forecasting horizons may have had a higher (lower) level of activity in terms of trading and volume. Another explanation could be that forecasts in the short run may be dominated by noise, with the noise being averaged out over the longer run, thus producing better longer-term forecasts. Mark (1995) and Chinn and Meese (1995) provide this rationale for similar evidence based on structural models. Further research in this area is needed.

APPENDIX A

In compliance with today's futures currency market, a more precise form of relationship (1) may be stated as

$$s_t = f\left(\sum_{j=0}^{N} p^M_{(t-k-j)}, \sum_{j=0}^{N} p^J_{(t-3k-j)}, \sum_{j=0}^{N} p^S_{(t-6k-j)}, \sum_{j=0}^{N} p^D_{(t-9k-j)}\right), \quad \text{(A.1)}$$

where

- s_t = first difference of the natural log of daily spot rate at time t, that is, $\ln S_t - \ln S_{t-1}$
- p^M = first difference of the natural log of futures price at time t for the contract maturing in March; that is, $\ln P^M_t - \ln P^M_{t-1}$
- p^J = first difference of the natural log of futures price at time t for the contract maturing in June, that is, $\ln P^J_t - \ln P^J_{t-1}$
- p^S = first difference of the natural log of futures price at time t for the contract maturing in September, that is, $\ln P^S_t - \ln P^S_{t-1}$
- p^D = first difference of the natural log of futures price at time t for the contract maturing in December, that is, $\ln P^D_t - \ln P^D_{t-1}$
- k = k days ahead horizon of the forecast
- N = a given number of days, which is not necessarily constant over the contracts. It represents a window that is used to include information that may be available in the market. The number of days in the window may vary according to the expiration of the contract. For example, a larger window may be used for a nearby contract under the premise that the closest price offers higher degree of certainty.

Further specification of the above relationship will depend upon the methodology employed at the estimation level. Additive linear and nonlinear forms as well as free forms are among the alternatives that are considered.

Based upon prior empirical work, relationship (A.1) is augmented to include, in addition to the *futures prices,* a brief history of past *spot prices;* that is

$$s_t = f\left(\sum_{j=0}^{N} p^M_{(t-k-j)}, \sum_{j=0}^{N} p^J_{(t-3k-j)}, \sum_{j=0}^{N} p^S_{(t-6k-j)}, \right.$$
$$\left. \sum_{j=0}^{N} p^D_{(t-9k-j)}, \sum_{j=0}^{N} s_{(t-k-j)}\right) \tag{A.2}$$

Because relationship (A.1) uses only *futures prices* (FP) on its right-hand side to forecast *spot rates,* this specification for simplicity is referred to as FP model in the text. Similarly, relationship (A.2), which uses *futures prices* (FP) and *spot prices* (SP) to forecast spot prices, is referred to as FPSP.

The performance of the above two models is gauged against a third model, the random-walk model (RW). The RW model assumes that the spot prices at time t are the best unbiased estimators of the future spot. The equation for the random walk is given as follows:

$$s_t = \varepsilon_t \tag{A.3}$$

Further refinements in the specifications of relationships (A.1) and (A.2) require extensive experiments embracing, simultaneously, three fronts: lag structure, data coverage, and estimation techniques. A brief summary of the outcomes of some of these experiments follows:

(a) Considering the broad scope and the dimensions of modeling/estimation effort in this article, that is, *three* models, *five* currencies, *three* techniques, and *ample* lag structures (hence specifications), employment of a set of highly elaborate sensitivity tests within the estimation and forecasting phases has to be scaled down. To this end, the lag structure at the estimation level is set somewhat a priori based on limited experiments with the magnitudes of the Hsiao (1981) final prediction error (FPE). These are later augmented by three measures as the guiding criteria for in-sample prediction: the Aikaki (1969) final prediction error (FPE), the Aikaki (1974) information criterion, and the Schwartz (1978) criterion. The max cap in these tests for inclusion of lagged variables on the right-hand side of relationships

(A.1) and (A.2) is set at 13 (for the FP model) and 17 (for the FPSP model). Also, to offer a forecast horizon of at least 1 month, the initial *length* of the lag is set at 30 days over a 4-day window, that is, -30, -31, -32, and -33 days for each of the independent variables (futures prices and spot rates, if any). The choice of a 4-day window is to avoid any possibility of a spike on a single day.

(b) The above process is further expanded when other additional parameters are involved. For example, for LWR limited preliminary tests yield the choice of ξ (the smoothing parameter). For the various values of ξ, see Table IV.

(c) In gauging the *out-of-sample* forecasting performance of the various alternatives that are estimated, MSE and RMSPE are the prime criteria. However, as noted above, this phase of the analysis is somewhat limited, considering the extensive number of options available, that is, combinations of models, techniques, forecasting horizons, and currencies. Resort to high-level statistical tests of significance *across* models is not feasible because of the extensive dimensions already considered here. Further extension employing the tests provided by Mizrach (1991) (see also Mizrach [1992, p. S160]) and Diebold and Mariano (1995) could constitute viable future research.

(d) For the NN model, the learning threshold for the network is set at 0.001. The initial rate and momentum are set at 0.1. The architecture is recurrent and the model is composed of one output layer, and inputs ranging from 2 to 18 variables, depending on the model specification chosen. The number of hidden layers is set at 10 arbitrarily.

(e) In addition to the above, the results of all the modeling efforts and techniques are offered for analysis and comparison. In general, the sensitivity analyses are limited mostly to the estimation phase in search of the best *in-sample* forecasting models; the estimated models/specifications are rarely changed because of their *out-of-sample* forecasting performance.

To recapitulate the above points and cast them in terms of the final forms of relationships (A.1) and (A.2), the following points are in order. The initial values of the parameters N and K in relationships (A.1) and (A.2) are set, respectively, at 4 days and 30 days. This yields a minimum of 30-day-ahead forecasts. To expand this horizon, especially for out-of-sample forecasts, a 30-day moving window is chosen. This window runs up to 180 days ahead in total. The out-of-sample forecasts are completely ex ante; they are made on the basis of information that is exogenously available beyond the current time period.

APPENDIX B

This appendix provides a very brief review of locally weighted regression (LWR) and artificial neural networks (NN).

Locally Weighted Regression (LWR)

LWR is an adaptation of the iterated weighted least-squares procedure and is used to smooth time-series plots by fitting a regression surface to data. That is, the dependent variable is smoothed as a function of the independent variables in a moving process, much like a moving average is computed for a time series. An early development of this technique was advanced by Cleveland (1979). Later, Cleveland and Devlin (1988) and Cleveland, Devlin, and Grosse (1988) improved the technique further.

The LWR model is represented by

$$y_i = g(x_i) + e_i \tag{B.1}$$

where $g(x_i)$ is the expected value of the regression surface, and e_i is a normal random variable with mean zero and variance σ^2. By fitting a linear function with the use of a weighted least-squares technique, the weights $w_i(x)$ are given by

$$w_i(x) = W\left(\frac{|x_i - x|}{d(x)}\right) \tag{B.2}$$

where x is any point in the space of the factors and $d(x)$ is the distance of x to the nth nearest neighbor of x_i.

In simpler terms, the fitted values are computed with the use of the nearest-neighbor technique, in which n number of nearest neighbors to x are selected and then weighted according to the distance of x from x_i. Like a nearest-neighbor estimator, LWR fits the surface at a point $x = x^*$ as a function of the y values corresponding to the k_T nearest neighbor of x^*. The model estimates $g(x^*)$ as the fitted value from the regression surface rather than as a simple average of the y values. This estimation would correspond to a simple average only in the very unlikely case that the constant term is the sole regressor with explanatory power.

To compute the LWR estimate of the surface at a point x^*, that is, $g(x^*)$, a smoothing constant (ξ) is defined such that $0 < \xi \leqslant 1$, and $k_T = \text{int}(\xi \cdot T)$, where int$(\cdot)$ rounds down to the nearest integer. The x_i's are ranked by Euclidean distance from x^* and are defined as $x_1^*, x_2^*, \ldots, x_T^*$. In this manner, x_1^* is closest to x^*, x_2^* is the second closest to x^*, and so on. The smoothing constant ξ determines the number of nearest neighbors used, and hence the degree of smoothing.

Artificial Neural Networks (NN)

The term *neural networks* describes a class of models that are also named in the literature as neural computation, artificial neural systems, and parallel distributed models. Caudill (1989) may explain it best: The artificial neural network (NN) represents a computing system that consists of interconnected processing units that process information according to the external inputs. Such models are designed to mimic the way the human brain acquires and organizes knowledge.

The neural model has two main characteristics: the input of signals from other units or the exterior of the model to a given unit, and the production of a signal by a neuron, which is passed on to other units. A neuron is a unit that receives activity and responds to other neurons if that activity is above a given threshold. In each case, every unit works independently of the other units. That is, for every two units connected to each other, a weight is assigned, which weighs the signal transmission accordingly. The total network behavior is therefore determined by the weight structures. If the network is in the learning phase, it constantly changes its weights, and hence modifies the knowledge represented by them.

There are two main classes of neural network architecture: feedforward methods, and feedback or recurrent nets. In the case of the feed-

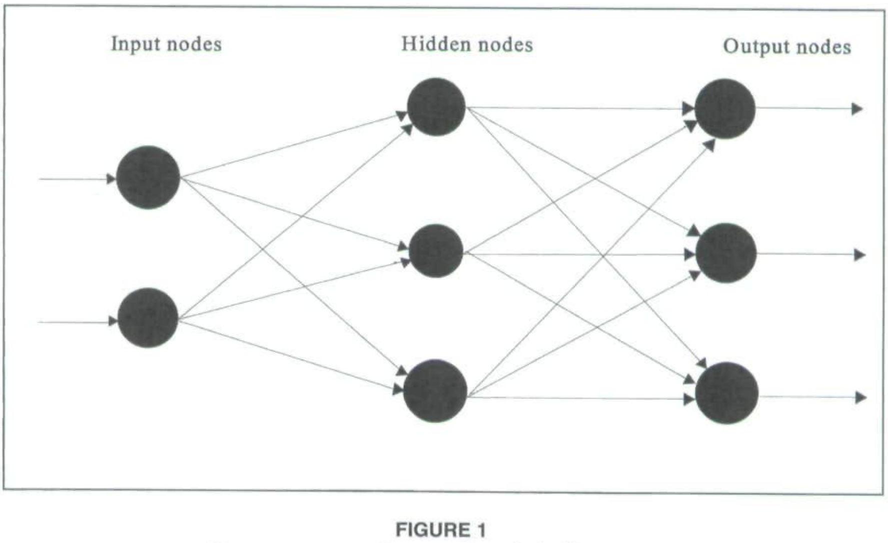

FIGURE 1
Recurrent neural network with hidden units.

forward networks, the input flows in only one direction to the output, with no feedback to previously active networks [see Hornik, et. al. (1989) for a mathematical treatment of this architecture]. The feedback or recurrent nets, on the other hand, have complete connectivity, with no distinction made between the input, the hidden units, and the output units. The input of the recurrent net at time *t* is therefore continuously modified by the previous output. That is, the weight modification between the nodes must be propagated all the way back to between the input layers and the hidden layers rather than only between the input nodes and the output nodes.

A three-layer recurrent network with hidden units resembles Figure 1.

BIBLIOGRAPHY

Akaike, H. (1969): "Filling Autoregressive Models for Predictions," *Annals of the Institute of Statistical Mathematics*, 21:243–247.

Akaike, H. (1974): "A New Look at Statistical Model Identification," *IEEE Transactions on Automatic Control 19*:716–723.

Backus, D. (1984): "Empirical Models of the Exchange Rate: Separating the Wheat from the Chaff," *Canadian Journal of Economics*, 17(4):824–846.

Baillie, R. T., and Bollerslev, T. (1989): "The Message in Daily Exchange Rates: A Conditional Variance Tale," *Journal of Business and Economics Statistics*, 7(3):297–305.

Bollerslev, T. (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69(3):542–547.

Boughton, J. M. (1987): "Tests of the Performance of Reduced-Form Exchange Rate Models," *Journal of International Economics*, 23($\frac{1}{2}$):41–56.

Boughton, J. M. (1988): "Exchange Rates and the Term Structure of Interest Rates," *International Monetary Fund Staff Papers*, 35(1):36–62.

Caudill, M. (1989): "Neural Networks Primer," *The Magazine of Artificial Intelligence in Practice*.

Chinn, M. D. (1991): "Some Linear and Nonlinear Thoughts on Exchange Rates," *Journal of International Money and Finance*, 10(2):214–230.

Chinn, M. D., and Frankel, J. (1994): "Patterns in Exchange Rate Forecasts for Twenty-Five Currencies," *Journal of Money, Credit and Banking*, 26(4):759–770.

Chinn, M. D., and Meese, R. A. (1995): "Banking on Currency Forecasts: How Predictable is Change in Money?" *Journal of International Economics*, 38:161–178.

Chowdhury, A. (1991): "Futures Market Efficiency: Evidence from Cointegration Tests," *Journal of Futures Markets*, 11(5):577–589.

Cleveland, W. S. and Devlin, S. J. (1988): "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association*, 83(403):596–610.

Cleveland, W. S. (1979): "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association,* 74(368):829–836.

Cleveland, W. S., Devlin, S., and Grosse, E. (1988): "Regression by Local Fitting: Methods, Properties, and Computational Algorithms," *Journal of Econometrics,* 37(1):87–114.

Diebold, F. X. (1988): *Empirical Modeling of Exchange Rate Dynamics.* New York: Springer-Verlag.

Diebold, F. X., and Nerlove, M. (1989): "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics,* 4(1):1–22.

Diebold, F. X., and Nason, J. A. (1990): "Nonparametric Exchange Rate Prediction?," *Journal of International Economics,* $28(\frac{3}{4})$:315–332.

Diebold, F. X., and Mariano, R. (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics,* 13(3):253–263.

Doukas, J., and Rahman, A. (1987): "Unit Roots Tests: Evidence from the Foreign Exchange Futures Market," *Journal of Financial and Quantitative Analysis,* 22(1):101–108.

Driskill, R. A., and Sheffrin, S. M. (1981): "On the Mark: Comment," *American Economic Review,* 71(5):1068–1074.

Engel, C. (1994): "Can the Markoff Switching Model Forecast Exchange Rates?," *Journal of International Economics,* $36(\frac{1}{22})$:151–165.

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

Finn, M. G. (1986): "Forecasting the Exchange Rate: A Monetary or Random Walk Phenomenon?" *Journal of International Money and Finance,* 5(2):181–193.

Frankel, J. A. (1982): "The Mystery of Multiplying Marks: A Modification of the Monetary Model," *Review of Economics and Statistics,* 64(3):515–519.

Frankel, J. A. (1984): "Tests of Monetary and Portfolio Balance Models of Exchange Rate Determination," in F. O. Bilson and R. C. Marston (Eds.), *Exchange Rate Theory and Practice,* Chicago: University of Chicago Press.

Friedman, J., and Vanderstell, S. (1982): "Short-Run Fluctuations in Foreign Exchange Rates: Evidence from the Data 1973–1979," *Journal of International Economics,* 13:171–186.

Granger, C. W. J. (1986): "Developments in the Study of Cointegrated Economic Variables," *Oxford Bulletin of Economics and Statistics,* 48:213–228.

Hakkio, C. S., and Rush, M. (1989): "Market Efficiency and Cointegration: An Application to the Sterling and Deutschemark Exchange Markets," *Journal of International Money and Finance,* 8:75–88.

Haynes, S., and Stone, J. (1981): "On the Mark: Comment," *American Economic Review,* 71(5):1060–1067.

Hoffman, D., and Schlagenhauf, D. E. (1983): "Rational Expectations and Monetary Models of Exchange Rate Determination: An Empirical Examination," *Journal of Monetary Economics,* 11(2):247–260.

Hooper, P., and Morton, J. (1982): "Fluctuations in the Dollar: A Model of Nominal and Real Exchange Rate Determination," *Journal of International Money and Finance,* 1(1):39–56.

Hornik, K., Stinchcombe, M. and White, H. (1989): "Multilayer Feedforward Networks Are Universal Approximators," *Neural Networks,* 2:359–366.

Hsieh, D. A. (1989): "Testing for Nonlinear Dependance in Daily Foreign Exchange Rates," *Journal of Business,* 62(3):339–368.

Kawaller, I., and Koch, T. K. (1984): "Cash and Carry Trading and the Pricing of Treasury Bill Futures," *Journal of Futures Markets,* 4:115–123.

Kuan, C.-M., and Liu, T. (1992): "Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks," Working Paper No. 92-0128, Department of Economics, University of Illinois, Urbana-Champaign.

MacDonald, S. S., and Hein, S. E. (1993): "An Empirical Evaluation of Treasury-Bill Futures Market Efficiency: Evidence from Forecast Efficiency Tests," *Journal of Futures Markets,* 13(2):199–211.

Mark, N. C. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *The American Economic Review,* 85(1):201–218.

McDonald, R., and Taylor, M. (1991): "Foreign Exchange Market Efficiency and Cointegration: Some Evidence from the Recent Float," *Economic Letters,* 29:63–68.

McGuirk, A., Robertson, J. and Spanos, A. (1993): "Modeling Exchange Rate Dynamics: Non-Linear Dependence and Thick Tails," *Econometric Reviews,* 12:33–63.

Meese, R. (1984): "The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?" in *Exchange Rates and International Macroeconomics,* ed. by Jacob Frenkel. Chicago: University of Chicago Press.

Meese, R., and Rogoff, K. (1983): "Empirical Exchange Rate Models of the Seventies: Do they Fit Out of Sample?" *Journal of International Economics,* 14($\frac{1}{2}$):3–24.

Meese, R., and Rose, A. K. (1991): "An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination," *Review of Economic Studies,* 58(195):603–619.

Mizrach, B. (1991): "Forecast Comparison in L_2," Working Paper, Department of Finance, Wharton School.

Mizrach, B. (1992): "Multivariate Nearest-Neighbour Forecasts of EMS Exchange Rates," *Journal of Applied Econometrics,* 7:S151–S163.

Mussa, M. (1979): "Empirical Regularities in the Behavior of Exchange Rates and Theories of the Foreign Exchange Market," in K. Brunner and A. Meltzer (Eds). *Policies for Employment, Prices, and Exchange Rates.* New York: North Holland, Carnegie-Rochester Conference, Vol. II.

Rasulo, J. A., and Wilford, D. S. (1980): "Estimating Monetary Models of the Balance of Payments and Exchange Rates: A Bias," *Southern Economic Journal,* 47(1):136–146.

Schinasi, G., and Swamy, P. A. (1989): "The Out-of-Sample Forecasting Performance of Exchange Rate Models when Coefficients Are Allowed to Change," *Journal of International Money and Finance,* 8:375–390.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics,* 6:461–464.

Schwarz, T. and Laatsch, F. (1991): "Dynamic Efficiency and Price Leadership in Stock I ndex Cash and Futures Markets," *Journal of Futures Markets,* 11(6):669–683.

Shen, C.-H., and Wang, L. R. (1990): "Examining the Validity of a Test of Futures Market Efficiency: A Comment," *Journal of Futures Markets,* 10(2):195–196.

Somanath, V. S. (1986): "Efficient Exchange Rate Forecasts: Lagged Models Better than the Random Walk?" *Journal of International Money and Finance,* 5(2):195–220.

Woo, W. T. (1985): "The Monetary Approach to Exchange Rate Determination under Rational Expectations," *Journal of International Economics,* 18:1–16.

Yadav, P., and Pope, P. (1991): "Testing Index Futures Market Efficiency Using Price Differences: A Critical Analysis," *Journal of Futures Markets,* 11(2):239–252.