



Multivariate Nearest-Neighbour Forecasts of EMS Exchange Rates

B. Mizrach

Journal of Applied Econometrics, Volume 7, Issue Supplement: Special Issue on Nonlinear Dynamics and Econometrics (Dec., 1992), S151-S163.

Stable URL:

<http://links.jstor.org/sici?sici=0883-7252%28199212%297%3CS151%3AMNFOEE%3E2.0.CO%3B2-I>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of Applied Econometrics is published by John Wiley & Sons, Ltd.. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/jwileytd.html>.

Journal of Applied Econometrics
©1992 John Wiley & Sons

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

MULTIVARIATE NEAREST-NEIGHBOUR FORECASTS OF EMS EXCHANGE RATES

B. MIZRACH

Federal Reserve Bank of New York, Research Function, 33 Liberty Street, New York, NY 10045, USA.

SUMMARY

Exchange rate modelling has been a persistent puzzle for international economists. Forecasts from popular models for the exchange rate generally fail to improve upon the random walk out-of-sample. While a multivariate nonparametric approach provides useful information about exchange rates, the model produces forecasts superior to the random walk for only one of the three EMS currencies examined. Using a statistic developed in Mizrach (1991), I find that the forecast improvement, a 4.5 per cent reduction in mean squared error for the Lira in daily returns, is not statistically significant. A cross-validation exercise suggests that the improvement is also not robust.

1. INTRODUCTION

Modelling exchange rates has been a particularly elusive task. There are numerous approaches in the literature, yet when Meese and Rogoff (1983) examined an exhaustive set of linear time-series, reduced-form and structural models,¹ they found no model that consistently produced forecasts superior to the random walk. Accounting for the poor performance of empirical exchange rate models has been on the research agenda of many economists.

An emerging body of work has emphasized nonlinearities as a possible key to the puzzle. The non-normality of exchange rate returns has motivated this research. The 'heavy-tailed' distributions in the log differences of exchange rates have been noted by a number of authors, including Westerfield (1977), Boothe and Glassman (1987) and Hsieh (1988). The generalized autoregressive conditional heteroscedasticity model of Engle (1982) and Bollerslev (1986), has been utilized extensively as an explanation of the leptokurtosis in asset returns.² Domowitz and Hakkio (1985), Diebold and Pauly (1988), Hsieh (1989a), and Diebold and Nerlove (1989) have all made contributions to the exchange rate literature using the GARCH model. Hsieh (1989b) has found evidence of nonlinear dependence after accounting for GARCH effects, using the test of Brock, Dechert and Scheinkman (1987). None of these papers has established that the nonlinearities may be useful for point prediction though.

A variety of nonlinear theoretical models have been suggested as explanations for the exchange rate data-generating process. As Mark (1985) has noted, if one relaxes the assumption that the representative agent is risk-neutral, the standard asset pricing model no

¹ Meese and Rogoff (1983) examine flexible-price and sticky-price monetary models, the sticky-price asset model, and several univariate time-series models.

² The GARCH models take no particular stance on the predictability of the conditional mean. It may be the case that the returns do follow a random walk, while the squared returns do not.

longer implies that prices will follow a random walk. Krugman (1988) has emphasized that target zones will make exchange rates depend nonlinearly on fundamentals. In the European Monetary System³ (EMS), informal target zones are in effect throughout much of my sample.

Recent econometric advances have facilitated empirical analysis of nonlinear models. Engel and Hamilton (1990) have applied Hamilton's (1989) model of regime switching to exchange rates. Meese and Rose (1990), Diebold and Nason (1990), and LeBaron (1990) have utilized nonparametric procedures. The nonparametric approaches allow the data analyst to consider any model with a smooth relationship between exchange rates and fundamentals.

In appraising the nonlinear analyses of exchange rates, one is surprised at the great difficulties very sophisticated models have in improving upon the random walk out-of-sample. Meese and Rose do not find any statistically important nonlinearities in the Bretton Woods fixed exchange rate regime. In Diebold and Nason's examination of 10 OECD currencies, the random walk has the lowest mean squared error in at least five of 10 cases for all the models they consider. LeBaron finds improvements of around 10 per cent in his low-volatility group for the West German Deutschmark and British pound,⁴ but not for the overall sample. Flood, Rose and Mathieson (1990) have tested a model similar to Krugman's. They uncover important nonlinearities in-sample for seven EMS/Deutschmark exchange rates, but again, the nonlinearity does not improve upon the random walk in out-of-sample forecasts.

This paper extends the nonlinear modelling of exchange rates to a multivariate setting. I attempt to incorporate structural information into a nonparametric analysis by looking at the managed system of exchange rates in the EMS. Since 1979, countries in the monetary system have attempted to coordinate policies so as to keep their exchange rates in alignment. I look at the three major EMS currencies, the French franc, Italian lira and the Deutschmark, at a daily frequency, throughout the entire floating exchange rate period. I estimate a multivariate nonlinear model for these currencies using a generalization of the nearest-neighbours procedure first proposed by Mack and Rosenblatt (1979).

I find that multivariate information is an important determinant of exchange rates. For all three currencies a multivariate model is superior to a univariate model both in and out-of-sample. The random walk is still a tenacious contender though. Only for the lira do I find a model that outperforms the random walk in forecasting out-of-sample.

The improvement for the lira is substantial, a 4.5 per cent reduction in mean squared prediction error in a series of over 750 daily returns. Given the magnitude of this improvement I do a variety of diagnostics. Using a statistic developed in Mizrach (1991), I find that the reduction in mean squared error is not statistically significant.

To test the sensitivity of my inference to data mining I try to cross-validate the forecast by reversing the estimation and forecast samples. If my result for the lira is more than a statistical artifact, the same model fitted to the forecast sample should predict as well as it did in the sample from which it was first estimated. I find, however, that the model that forecasts best into the late 1980s, forecasts very poorly backwards into the 1970s.

Section 2 develops the univariate nearest-neighbours model. This section reinforces the conclusions of Meese and Rose, and Diebold and Nason, on weekly data with a larger daily data sample. No serious competitor to the random walk emerges. In Section 3 I discuss the EMS and why a multivariate approach might be useful. Section 4 develops and applies the

³ An excellent source of institutional information is the paper by Ungerer *et al.* (1986). See section 3 for further discussion.

⁴ Even these modest forecast improvements are open to question. None of these forecast improvements has been shown to be statistically significant, nor have the results been cross-validated.

multivariate estimator. Section 5 evaluates statistically the forecasts, and Section 6 then tries to cross validate them. Section 7 concludes with some thoughts about empirical exchange rate modelling in light of the paper's results.

2. THE UNIVARIATE NEAREST-NEIGHBOURS ESTIMATOR

This section develops the uniformly weighted nearest-neighbours estimator. In part 1 consistency of the estimator is established under fairly general regularity conditions. Part 2 considers a way to weight the data using least-squares regressions. Univariate empirical results are presented in part 3.

Neighbour Selection and Consistency

Consider the regression model:

$$y_t = f(x_t) + \varepsilon_t \quad t = 1, \dots, n \tag{1}$$

where $x_t = (x_{1t}, \dots, x_{pt})$ is a $1 \times p$ vector of explanatory variables, f is a smooth function mapping $R^p \rightarrow R$, and ε_t is an independent and identically distributed mean zero disturbance term. The nearest-neighbour regression function estimate at time $t \in [1, \dots, n]$ is given by:

$$\hat{f}(x_t) = \sum_{i=1}^n w_{it} I[\|x_i - x_t\| < \eta] y_i \tag{2}$$

where I is the indicator function, $\|\cdot\|$ is the Euclidean norm

$$d(x_i - x_t) = \left[\sum_{l=1}^p (x_{il} - x_{tl})^2 \right]^{0.5}, \tag{3}$$

η is some constant, and $w_{it} = (w_{1t}, \dots, w_{nt})$ is a sequence of probability weights.

Stone (1977) cleverly formulated the problem of consistent estimation through regularity conditions on weights for the neighbours. For example, let q be the cardinality of the set of nearest neighbours less than η from x_t , $\# \{k^{\eta}t\} = (k_{1t}, \dots, k_{qt})$. A simple average of the q neighbours would set $w_{it} = 1/q$ for all i such that $I[\|x_i - x_t\| < \eta]$, and 0 otherwise.

Stone proved consistency for probability weights satisfying the following necessary and sufficient conditions:

$$E \sum_{i=1}^n w_{it} f(x_i) \leq C E f(x_t) \quad \forall t \geq 1, \tag{4}$$

where C is some constant;

$$\text{Prob.} \left[\sum_{i=1}^n w_{it} I[\|x_i - x_t\| > a] \right] \rightarrow 0, \quad \forall a > 0; \tag{5}$$

$$\text{Prob.} \left[\max_i w_{it} \right] \rightarrow 0; \tag{6}$$

If the weights are uniform, quadratic or triangular, these revert to the familiar conditions that as $n \rightarrow \infty$, $q \rightarrow \infty$, but $q/n \rightarrow 0$. As the sample grows large the number of neighbours must go off to infinity, but at a slower rate than the sample size increases. Consistency becomes a matter of imposing a selection rule involving η . As a practical matter, our investigation will look over a range of q 's.

Locally Weighted Regression

Having selected the neighbours there are a variety of ways to fit the regression surface. Following Meese and Rose, and Diebold and Nason, I develop in this part the locally weighted regression procedure (LWR) introduced by Cleveland (1979). In LWR the conditional mean is determined from a least-squares regression on the neighbours.

Denote the $n \times q + 1$ matrix of neighbours by

$$K_n = \begin{bmatrix} 1 & k_{11} & \dots & k_{q1} \\ 1 & k_{12} & \dots & k_{q2} \\ \vdots & \vdots & & \vdots \\ 1 & k_{1n} & \dots & k_{qn} \end{bmatrix}, \quad (7)$$

which includes a constant term in the first column. The locally weighted regression estimate is

$$\hat{f}(x_t) = k_t (K_n' K_n)^{-1} K_n' Y_n, \quad (8)$$

where k_t is the t th row of K_n , and $Y_n = [y_1, \dots, y_n]'$. The least-squares estimates of the weights will correspond to a simple average, $1/q$, only in the unlikely event that the constant is the only significant regressor in (8). This procedure allows a smoothing of weights on the neighbours so as to minimize mean-square error in-sample.

A further downweighting of the distant neighbours proved helpful in the empirical analysis. Each of the q -nearest neighbours was weighted by Euclidean distance from the current state. I set $w_{it} = 1 - u$, where

$$u \equiv \frac{\|k_{it} - x_t\|}{\sum_{i=1}^q \|k_{it} - x_t\|}. \quad (9)$$

I also tried the tricube function, $w_{it} = (1 - u^3)^3$, suggested by Cleveland, but (9) proved slightly superior empirically.

Univariate Estimation and Forecasting

In all the empirical work that follows, the models were estimated on data for the period 2 January 1974–31 December 1985, a sample of 3063 daily observations. Regression functions were fitted over two subsamples, January 1974–12 March, 1979 and 13 March, 1979 to 31 December 1985. 13 March, 1979 coincides with the formation of the European monetary union. Out-of-sample forecasts were then computed for the period 1 January 1986 to 31 December 1988, a span of 757 observations. The data are spot bid daily closes from the London market for the French franc, Italian lira, and West German Deutschmark. They are expressed in currency unit per US dollar. The analysis was done with log differences to avoid any problems with stationarity.

Uniform Weights

I first utilize the uniformly weighted nearest-neighbour estimator. I specify autoregressive models for the conditional expectation of (1). x_t is a vector of p -lags of the spot exchange rate. In the multivariate analysis of section 3, x_t will be a $(p \times 3)$ vector of lags of all three exchange rates. The neighbour selection rule (2) will choose spot exchange rates for which the distance

(3) is small. These will be lags of the exchange rate, but they may come from any portion of the estimation sample.

As a baseline nonparametric estimator I looked at an AR(1) model with q neighbours. The weights in this section are simple averages, $w_i = 1/q$. Consider the first two columns of Table I for each currency. # kNN is the number of neighbours. Lags indicates the dimension of the autoregression. In this baseline table, p is always equal to one. I include the random walk as a benchmark.

The next two columns are in the in-sample standard errors for the two regimes of the estimation sample. The period from January 1974 to 12 March 1979 is the pre-EMS sample, and the second is the EMS policy regime. The period January 1986 to December 1988 is reserved for forecasting.

While the standard errors diminish with the number of neighbours, the random walk easily tops a model with 90 neighbours, in-sample, for all three currencies. The forecasts are, in percentage terms, even poorer. The one-neighbour model's mean squared prediction error is double the random walk for the lira, and nearly double for the franc and Deutschmark. While adding more neighbours again moves us closer to the random walk, the approach is quite slow. Though not reported in the table, models with as many as 500 neighbours did not encroach upon the random walk. Clearly more work needs to be done.

Table I. Uniformly Weighted Nearest Neighbour Estimator

Model		Standard errors*		MSPE†
# kNN	Lags	Jan. 1974–12 Mar. 1979	13 Mar. 1979–Dec. 1985	Jan. 1986–Dec. 1988
<i>(a) French franc</i>				
1	1	7.208	10.415	9.291
5	1	7.032	10.317	9.287
30	1	6.870	9.991	8.739
60	1	6.916	10.003	8.675
90	1	6.860	9.904	8.540
RW		5.153	7.324	4.830
<i>(b) Italian lira</i>				
1	1	6.721	9.773	9.766
5	1	6.631	9.664	9.756
30	1	6.202	9.304	9.239
60	1	6.171	9.284	9.118
90	1	6.105	9.182	8.933
RW		5.152	6.993	4.830
<i>(c) West German Deutschmark</i>				
1	1	7.394	9.984	10.130
5	1	7.305	9.963	10.089
30	1	7.160	9.740	9.559
60	1	7.231	9.784	9.511
90	1	7.193	9.704	9.368
RW		5.256	7.086	5.159

* Standard error = $(e'e/(n-k))^{0.5}$. All numbers are $\times 10^{-3}$.

† Mean squared prediction error = All numbers are $\times 10^{-5}$.

kNN is the number of nearest neighbours used. Lags refer to the order of the model. 1 indicates an AR(1) model. RW is the random walk.

Local Regression

I next turn to local regression as a way to improve upon the uniform weights. Results for unweighted regressions like (8) for the three currencies are in Table II(a)–(c). In all three cases there is dramatic improvement. The nonparametric fit now easily tops the random walk in-sample. Still none of these models performs more capably than the random walk out-of-sample. One feature that does seem to merit attention is that the one-neighbour forecasts are among the best. Choosing a model on the basis of in-sample fit would not be a good strategy.

My next step is to use a weighted regression procedure. Using the local distance measure (9), I assign greater weight *a priori* to close neighbours. Downweighting the outlying neighbours improved the forecasts over Table II, but the gains are somewhat illusory. The one-neighbour models provide the best forecasts; the weights simply make the model with more neighbours perform like the one-neighbour models.

I also look at higher order autoregressions in Table III. A model with three lags and one nearest neighbour offers the most striking improvement over Tables I and II. This model has the lowest mean squared prediction error for all three currencies. In the case of the lira, the model is the first to outperform the random walk.

To ensure that my inference was not somehow being contaminated by just a few outliers, I also look at the mean absolute forecast errors. The results are qualitatively the same using

Table II. Unweighted local regression

Model		Standard errors*		MSPE†
# <i>kNN</i>	Lags	Jan. 1974–12 Mar. 1979	13 Mar. 1979–Dec. 1985	Jan. 1986–Dec. 1988
<i>(a) French franc</i>				
1	1	5·152	7·315	4·878
5	1	5·148	7·320	4·872
30	1	5·033	7·284	5·033
60	1	5·003	7·288	5·223
90	1	4·955	7·281	5·516
RW				4·830
<i>(b) Italian lira</i>				
1	1	5·104	6·981	4·857
5	1	5·031	6·943	4·847
30	1	4·409	6·868	5·044
60	1	4·160	6·727	5·755
90	1	4·042	6·658	6·207
RW				4·814
<i>(c) West German Deutschmark</i>				
1	1	5·249	7·087	5·176
5	1	5·255	7·073	5·250
30	1	5·220	7·073	5·399
60	1	5·165	7·044	5·513
90	1	5·127	6·996	6·260
RW				5·159

* All numbers are $\times 10^{-3}$.

† All numbers are $\times 10^{-5}$.

#*kNN* is the number of nearest neighbours used. Lags refer to the order of the model. 1 indicates an AR(1) model. RW is the random walk.

Table III. Locally weighted nearest-neighbour regression forecasts 1 January 1986 to 31 December 1988

Model		Franc		Lira		Deutschmark	
# <i>kNN</i>	Lags	MSPE*	MAE†	MSPE*	MAE†	MSPE*	MAE†
1	1W	4·877	5·074	4·857	5·270	5·176	5·317
1	3W	4·894	5·076	4·805	5·244	5·177	5·315
5	1W	4·869	5·078	4·882	5·281	5·217	5·328
5	3W	4·900	5·100	4·880	5·276	5·342	5·317
RW		4·830	5·030	4·814	5·221	5·159	5·298

* All numbers are $\times 10^{-5}$.

† Mean absolute error = $|\sum_n (y - y^e)|/n$. All numbers are $\times 10^{-3}$.

kNN is the number of nearest neighbours used. Lags refer to the order of the model. 3 indicates an AR(1) model. RW is the random walk. The weights are given by (9) in the text.

this metric. The random walk is again superior to nonparametric models for the franc and the Deutschmark.

3. THE EUROPEAN MONETARY SYSTEM

Section 3 has two parts. The first briefly outlines some institutional detail about the EMS. Part 2 motivates why the EMS might lead to multivariate linkages in US dollar exchange rates.

The Exchange Rate Mechanism (ERM)

Since the European Monetary System was formed in March 1979, the central banks of member nations have attempted to coordinate their monetary policies. Belgium–Luxemburg, Denmark, France, Germany, Ireland, Italy and the Netherlands participate in the exchange rate mechanism (ERM).⁵ Currencies are allowed to fluctuate within a 2·25 per cent range, with the exception of Italy, which floats in a 6 per cent range.

The policy coordination among member nations has been successful. There have been seven major realignments since the system was formed, but the period has been one of unprecedented stability.⁶ A large number of academic studies⁷ have concluded that the EMS has succeeded in decreasing the volatility of the members' exchange rates.

Multinational Exchange Rate Links

Consider once again our canonical model (1) for the exchange rate. Let y_t be the spot exchange rate and x_t be a vector of fundamentals, (1) then nests a large number of models for the exchange rate, including all of those considered by Meese and Rose (1991). Relative money supplies are a fundamental component of virtually every model of exchange rates; for expository purposes assume that monetary policy is the stabilization tool.

⁵ The UK was the noticeable exception during my sample period.

⁶ The two years between March 1983 and July 1985 passed without any official revaluations. See Ungerer *et al.* (1986) for further discussion.

⁷ See Bodnar (1989), *inter alia*.

Suppose that the target zone in question is the franc/Deutschmark exchange rate. An expansion in the money supply of Germany to defend the franc would not necessarily imply a change in the US dollar/franc exchange rate. *Ceteris paribus*, though, this would imply an expansion of the German money supply relative to the US, and depreciate the Deutschmark relative to the dollar. A univariate model for the franc might miss the effect of intra-EMS intervention, but in a multivariate framework, the vector of dollar exchange rates *would* be affected, regardless of whether the intervention came from France or Germany.

This naturally suggests that the limitation of previous studies to single currency regression functions is unduly restrictive. I seek to model this below by generalizing the nearest-neighbours approach to a multivariate setting.

4. MULTIVARIATE ANALYSIS

In this sections I attempt to use the cross-currency information in forming the nonparametric regression function. I first develop the multivariate estimator and then again turn to a variety of different weighting functions.

The multivariate generalization is straightforward but notationally a bit cumbersome. Re-write (1) as a z -variate model

$$y_{jt} = f(x_{jt}) + \varepsilon_{jt} \quad j = 1, \dots, z \quad t = 1, \dots, n \quad (10)$$

A regression function for the j th model at time t is

$$\hat{f}(x_{jt}) = \sum_{j=1}^z \sum_{i=1}^n w_{jit} I[\|x_{ji} - x_{jt}\| < \eta] y_{ji}. \quad (11)$$

The norm is now the multivariate Euclidean norm

$$d(x_{ji} - x_{jt}) = \left[\sum_{j=1}^z \sum_{l=1}^p (x_{jli} - x_{jlt})^2 \right]^{0.5}. \quad (12)$$

The neighbour selection rule (11) says include all $z = 3$ exchange rates when the cumulative distance of the explanatory variables from their neighbours is less than some η . $k^{\eta t}$ is now a $q \times z$ matrix of neighbouring exchange rates at time t , and w_{jit} will be an $n \times z$ matrix. In the discussion that follows, I will refer to the neighbour selection distance function (3) as the univariate norm and to (12) as the multivariate norm. If K_n , from (8), is the $n \times q * z + 1$ matrix of neighbours for all three currencies, I will call the regression function multivariate.

Multivariate Empirical Results

I repeated the estimation and forecasting exercises of Section 2. The data and sample periods are identical, but I now undertake a multivariate analysis.

I found that the best forecasts came from using a multivariate local regression with neighbours selected by the univariate norm (3). Table IV reports results for these regressions.

A multivariate nearest-neighbour model for the lira with three lags and local weights provides the best forecast overall, a 4.5 per cent improvement over the random walk. While forecasts for the franc and Deutschmark improve substantially using neighbours from all three currencies, none outperforms the random walk.

I also looked at multi-step-ahead prediction. I forecast five steps ahead, using the same models as in Table IV. The results (not reported) for these weekly returns are no better than they were for the daily returns. The random walk wins the competition again, even for the lira.

Table IV. Multivariate nearest-neighbour regression univariate norm* forecasts 1 January 1986 to 31 December 1988

Model		Franc		Lira		Deutschmark	
# <i>kNN</i>	Lags	MAE †	MSPE ‡	MAE †	MSPE ‡	MAE †	MSPE ‡
1	1	5.105	4.848 (-0.28)	5.155	4.682 (-0.58)	5.352	5.228 (1.56)
5	1	5.189	5.145 (1.37)	5.232	4.867 (0.13)	5.432	5.515 (2.06)
1	1W	5.097	4.838 (1.12)	5.158	4.693 (-0.54)	5.352	5.224 (1.55)
1	3W	5.095	4.843 (0.21)	5.110	4.594 (-0.94)	5.362	5.244 (1.62)
5	1W	5.147	4.963 (1.52)	5.170	4.723 (-0.41)	5.389	5.291 (1.86)
5	3W	5.317	5.239 (3.29)	5.203	4.732 (-0.37)	5.475	5.499 (2.51)
RW		5.030	4.830	5.221	4.814	5.298	5.159

* # *kNN* is the number of nearest neighbours used. Lags refer to the order of the model. 1 indicates an AR(1) model. Models with local weights have a W. RW is the random walk. M-statistics (in parentheses) are given by equation (14) in the text. It has an asymptotic standard normal distribution.
 † All numbers are $\times 10^{-3}$.
 ‡ All numbers are $\times 10^{-5}$.

Despite my success with the lira, two questions remain open. Is a 4.5 per cent improvement large in a statistical sense. Secondly, is this model simply an artefact of the data, uncovered by an exhaustive data search. Section 5 tries to answer the former and Section 6 the latter.

5. STATISTICS FOR FORECAST COMPARISON

While the improvement for the lira does seem dramatic, there is no assurance that this improvement is statistically significant. Below, I develop a statistic for assessing forecast improvement under weak population assumptions. In Section 6 I will try to see if the improvement is robust through a cross-validation exercise.

Consider two forecasts, \tilde{y}_1 and \tilde{y}_2 , and let the respective forecast errors be $e_1 \equiv y - \tilde{y}_1$ and $e_2 \equiv y - \tilde{y}_2$. Let $MSPE_i$ be the mean squared prediction error of forecast *i*:

$$MSPE_i \equiv 1/n \sum_{t=1}^n e_{it}^2 \tag{13}$$

One would like some way to determine if two mean squared prediction errors are statistically different from one another. The standard *F*-test is tempting but not appropriate here. Even if the errors were normal the two MSPEs are not draws from independent random samples.

If the forecast errors were from a mean zero bivariate normal population (E_1, E_2), with correlation ρ_E and standard deviations, σ_1 and σ_2 , a straightforward test of the forecast improvement is available. Granger and Newbold (1986) and Meese and Rogoff (1983) noticed an easy way to test the equality of sample MPSEs through the following transformation. Let $U = E_1 - E_2$, and $V = E_1 + E_2$. Then, (U, V) has a bivariate normal distribution with parameters $E[U] = \mu_1 - \mu_2 = 0$, $E[V] = \mu_1 + \mu_2 = 0$, $\text{var}(U) \equiv \sigma_U^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_E\sigma_1\sigma_2$, $\text{var}(V) \equiv \sigma_V^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_E\sigma_1\sigma_2$, and $\text{cov}(U, V) \equiv \sigma_{UV} = \rho\sigma_U\sigma_V$. In terms of the original

population, $\sigma_{UV} = \sigma_1^2 - \sigma_2^2$. If the mean squared prediction errors in the original population are equal, then the covariance in the transformed population must be zero. Given normality, a standard test of lack of correlation between U and V would be the most powerful test.⁸

Granger and Newbold's approach, unfortunately, cannot handle properly our exchange rate population. As I noted in the introduction, the non-normality of the returns is a critical stylized fact, and this property carries over to the forecast errors, e_1 and e_2 . I examined the errors from our top performer, the lira model with one neighbour, three lags and local weights. In the sample of 757 errors the mean error is 7.931×10^{-4} , and the excess kurtosis is 1.434. I can reject at the 99 per cent level that the forecast errors are unbiased, and at the 99.9 per cent level that the excess kurtosis is zero. The assumptions of unbiasedness and normality are clearly violated in the population. Given that we are working with time-series data, the forecast errors are also likely to be serially correlated.

Some recent work in Mizrach (1991) has extended the Granger and Newbold procedure to allow for biased errors, non-normality and heteroscedasticity. I refer the reader to that manuscript for further details. I show that the statistic

$$\frac{1/n \sum_{j=1}^n u_j v_j}{\left[1/n \sum_{t=-k}^k (1 - (|t|/k + 1)) S_{UVUV}(t) \right]^{1/2}}, \quad (14)$$

where

$$S_{UVUV}(t) = 1/n \sum_{j=1}^{n-|t|} u_j v_j u_{j+|t|} v_{j+|t|}, \quad (15)$$

is distributed asymptotically $N(0, 1)$ when the order of dependence is known to be k .

I compare the forecasts from Section 4 to the random walk in Table IV using the statistic (14). The inference one arrives at is counter to the conventional wisdom. Under the Granger and Newbold assumptions one would conclude that several models for the lira are better than the random walk.⁹ With the robust statistic (14) that conclusion is no longer substantiated.¹⁰ Our top predictor has a p-value of only 0.3472.

A second hurdle still remains for forecasters of the exchange rate. Even our limited success with the lira might be fragile because of data mining. The cross-validation exercise in Section 6 will lend support to this view.

6. CROSS-VALIDATION OF FORECASTS

There is a variety of judgemental inputs that enter the regression function. First, one must choose the number of nearest neighbours. Next, one must choose a univariate or vector autoregression, and then the number of lags to include in the specification. Apart from the

⁸ Fisher (1915) has derived the exact finite sample distribution of r , the sample correlation coefficient. In the case, where $\rho = 0$, this distribution simplifies to an incomplete beta. The statistic $\tau = r\sqrt{n-2}/\sqrt{1-r^2}$ has the Student t -distribution, with $n-2$ degrees of freedom.

⁹ The τ -statistics (see footnote 8) for the three leading models are -1.99 , -2.10 , and -3.28 . The first two are significant at the 95 per cent level, and the latter at the 1 per cent level.

¹⁰ Monte-Carlo results in Mizrach (1991) help explain these results. I find that the Granger and Newbold statistic is poorly sized in leptokurtic populations. One rejects the null hypothesis that the MSPEs are equal from three to four times too often in a 5 per cent test.

choice of a multivariate model, these choices were largely data-driven. I searched across permutations of these judgemental inputs, with forecast performance as the criterion.

A procedure recommended in the nonparametric literature is to 'allow the data to choose' these judgemental inputs. A method by which one might consistently specify the model is known as cross-validation.¹¹ As Leamer (1983) notes, in a criticism of cross-validation, the intent of cross-validation procedures is to avoid finding a model that fits well only for a given data set. One usually finds such a model by a specification search, and validation procedures try to perturb the data set in some way so as to see how robust the inference for the estimated model is. Below, I fit the regression function to the data used previously for forecasting.

The most straightforward cross-validation inference is using a split sample analysis. The first step is to partition the data into two samples, n_1 , and n_2 , leaving some portion of the sample uncontaminated by specification searches. Efron (1983) calls the sample covering the period March 1979–December 1985 our 'training' set. In the training set, one can data mine to any degree, secure that the other portion of the sample is available for *ex-post* comparisons.

I took the best model for the Italian lira, as determined by mean-squared prediction error for the period 1 January 1986–31 December 1988, the univariate norm, multivariate model with three lags (Table IV, line 4). I then repeated the nearest-neighbour estimation, using data from the forecast period, and then forecasting out-of-sample into 31 December, 1985–13 March, 1979. This model (and the one with only a single lag) no longer outperformed the random walk, as can be seen in Table V.

The split sample procedure does have its drawbacks. It overemphasizes stability of the underlying coefficients, even if, as in a nonparametric analysis, one is not interested in estimating them directly. Flood, Rose and Mathieson (1990) suggest that there may be as many as 12 regimes in our sample.¹² It is perhaps not surprising, then, that our inference was so fragile. Still, the failure to cross-validate the forecast for the lira is evidence that the victory over the random walk is a statistical artefact.

Table V. Cross-validation
of forecasts

# kNN	Italian lira	
	Lags	MSPE*
1	1W	3·936
1	3W	3·917
RW		3·819

*All MSPEs are $\times 10^{-5}$.
Models with local weights have a W. The forecast (backcast) interval here is 31 December 1985–13 March 1979.

¹¹ A standard early reference is the paper by Lachenbruch and Mickey (1968). Other important contributions include Stone (1974), Geisser (1975), and Wahba and Wold (1975), and Li (1984).

¹² These regimes correspond to the 11 realignments in the EMS between March 1979 and December 1989. Our forecast sample coincides with one of the longest stable periods. Between 24 January 1987 and 31 December 1989, 770 days passed without a realignment.

7. SUMMARY

From a theoretical perspective one might be quite surprised to find much empirical support for the random walk. To generate the random walk implication, strong restrictions on preferences and central bank intervention are needed. These are clearly violated in the real world, yet little persuasive empirical evidence has emerged that refutes the random walk hypothesis. This paper offers little evidence to the contrary. Several models for the lira produced substantial improvements in forecast performance, but none was statistically significant.

The fragility of the inference suggests that we are still a very long way from understanding the role that nonlinearities play. The paper does not present a structural model. In light of the cross-validation exercise, the forecast improvement must be viewed largely as a statistical artifact. The time-series representation for the 1980s is clearly different from the 1970s. Some deeper structure may still be lurking in the data, but until such a model is found, the random walk should remain a leading characterization of the exchange rate data-generating process.

ACKNOWLEDGEMENTS

I would like to thank Fabio Canova, Don Cox, Kathryn Dominguez, Frank Diebold, Tim Erickson, Karen Lewis, Bob Murphy, Lars Svensson and three anonymous referees for helpful comments and discussion, and Mark Coleman of Data Resources, Inc. for providing the data used in this analysis. Earlier versions of this paper were presented at the 1991 North American Winter Meetings of the Econometric Society in Washington, DC, and the UCLA Conference on Nonlinear Dynamics and Econometrics, 5–6 April, 1991. This paper was written while I was visiting at the Wharton School. The views expressed do not reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System. All errors are my own.

REFERENCES

- Bodnar, G. (1989), 'Target zones and Euro rates: A model of Eurocurrency interest rate differentials in the EMS'. Manuscript, Princeton University.
- Bollerslev, T. P. (1986), 'Generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics*, **31**, 307–27.
- Boothe, P., and D. Glassman (1987), 'The statistical distribution of exchange rates', *Journal of International Economics*, **22**, 297–319.
- Brock, W., W. D. Dechert, and J. A. Scheinkman (1987), 'A test for independence based on the correlation dimension'. SSRI Working Paper 8702, Department of Economics, University of Wisconsin.
- Cleveland, W. S. (1979), 'Robust locally weighted regression and smoothing scatterplots', *Journal of the American Statistical Association*, **74**, 829–836.
- Diebold, F., and J. Nason (1990), 'Nonparametric exchange rate prediction?', *Journal of International Economics*, **28**, 315–332.
- Diebold, F., and M. Nerlove (1989), 'The dynamics of exchange rate volatility: a multivariate latent factor ARCH model', *Journal of Applied Econometrics*, **4**, 1–22.
- Diebold, F., and P. Pauly (1988), 'Endogenous risk in a rational expectations portfolio balance model of the Deustchemark/dollar rate', *European Economic Review*, **32**, 27–54.
- Domowitz, I., and C. Hakkio (1985), 'Conditional variance and the risk premium in the foreign exchange market'. *Journal of International Economics*, **19**, 47–66.
- Efron, B. (1983), 'Estimating the error rate of a prediction rule. Improvement on cross-validation', *Journal of the American Statistical Association*, **78**, 316–331.
- Engel, C., and J. D. Hamilton (1990), 'Long swings in the dollar: are they in the data and do markets know it?', *American Economic Review*, **80**, 689–713.
- Engle, R. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation', *Econometrica*, **50**, 987–1008.

- Fisher, R., (1915), 'Frequency distribution of the values of correlation coefficients in samples from an indefinitely large population', *Biometrika*, **10**, 507–21.
- Flood, R. P., A. K. Rose, and D. J. Mathieson (1990), 'An empirical exploration of exchange rate target zones', NBER Working Paper 3543.
- Geisser, S. (1975), 'The predictive sample reuse method with applications'. *Journal of the American Statistical Association*, **70**, 320–328.
- Granger, C. W. J., and P. Newbold (1986), *Forecasting Economic Time Series*, Academic Press, San Diego.
- Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle'. *Econometrica*, **57**, 357–384.
- Hsieh, D. (1988), 'The statistical properties of daily foreign exchange rates: 1974–83', *Journal of International Economics*, **24**, 129–145.
- Hsieh, D. A. (1989a), 'Modeling conditional heteroscedasticity in daily foreign exchange rates', *Journal of Business Economics and Statistics*, **7**, 307–317.
- Hsieh, D. A. (1989b), 'Testing for nonlinear dependence in foreign exchange rates', *Journal of Business*, **62**, 339–368.
- Krugman, P. (1988), 'Target zones and exchange rate dynamics', NBER Working Paper 2481.
- Lachenbruch, P., and M. Mickey (1968), 'Estimation of error rates in discriminant analysis', *Technometrics*, **10**, 1–11.
- Leamer, E. (1983), 'Model choice and specification analysis,' in Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, Vol. I. North-Holland, Amsterdam, pp. 285–330.
- LeBaron, B. (1990), 'Forecast improvements using a volatility index', Working Paper, U. Wisconsin-Madison.
- Li, K. C. (1984), 'Consistency for cross-validated nearest neighbor estimates in nonparametric regression', *Annals of Statistics*, **12**, 230–240.
- Mack, Y. P., and M. Rosenblatt (1979), 'Multivariate k -nearest neighbor density estimates', *Journal of Multivariate Analysis*, **9**, 1–15.
- Mark, N. (1985), 'On time varying risk premia in foreign exchange markets: an econometric analysis', *Journal of Monetary Economics*, **16**, 3–18.
- Meese, R. A., and K. Rogoff (1983), 'Empirical exchange rate models of the seventies: do they fit out of sample?', *Journal of International Economics*, **14**, 3–24.
- Meese, R. A., and A. K. Rose (1990), 'Nonlinear, nonparametric, nonessential exchange rate estimation', *American Economic Review Papers and Proceedings*, **80**, 192–196.
- Meese, R. A., and A. K. Rose (1991), 'An empirical assessment of nonlinearities in models of exchange rate determination', *Review of Economics Studies*, **80**, 603–619.
- Mizrach, B. (1991), 'Forecast comparison in L_2 ', Working Paper, Department of Finance, Wharton School.
- Stone, C. J. (1974), 'Cross validity choice and assessment of statistical predictions', *Journal of the Royal Statistical Society, Series B*, **36**, 111–147.
- Stone, C. J. (1977), 'Consistent non-parametric regression', *Annals of Statistics*, **5**, 595–645.
- Ungerer, H., O. Evans, T. Mayer, and P. Young (1986), 'The European monetary system: recent development'. IMF Occasional Paper 48.
- Wahba, G., and S. Wold (1975), 'A completely automatic French curve: fitting spline functions by cross-validation', *Communications in Statistics*, **4**, 1–17.
- Westerfield, J. A. (1977), 'An examination of foreign exchange risk under fixed and floating rate regimes', *Journal of International Economics*, **7**, 181–200.