

EMPIRICAL EXCHANGE RATE MODELS OF THE SEVENTIES

Do they fit out of sample?

Richard A. MEESE*

University of California at Berkeley, Berkeley, CA 94720, USA

Kenneth ROGOFF

Board of Governors of the Federal Reserve System, Washington, DC 20551, USA

Received July 1981, revised version received April 1982

This study compares the out-of-sample forecasting accuracy of various structural and time series exchange rate models. We find that a random walk model performs as well as any estimated model at one to twelve month horizons for the dollar/pound, dollar/mark, dollar/yen and trade-weighted dollar exchange rates. The candidate structural models include the flexible-price (Frenkel–Bilson) and sticky-price (Dornbusch–Frankel) monetary models, and a sticky-price model which incorporates the current account (Hooper–Morton). The structural models perform poorly despite the fact that we base their forecasts on actual realized values of future explanatory variables.

1. Introduction

This study compares time series and structural models of exchange rates on the basis of their out-of-sample forecasting accuracy. We find that a random walk model would have predicted major-country exchange rates during the recent floating-rate period as well as any of our candidate models.¹ Significantly, the structural models fail to improve on the random walk model in spite of the fact that we base their forecasts on actual realized values of future explanatory variables.

*Both authors were at the Federal Reserve Board when this paper was written. This paper is a revised version of a paper presented at the International Monetary Fund and at the December 1981 Meetings of the Econometric Society. Robert Flood, Jeffrey Frankel, Robert Hodrick, Peter Hooper, and Julio Rotemberg gave us helpful comments on an earlier draft. We are indebted to Julie Withers and Catherine Crosby for research assistance. This paper represents the views of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System.

¹Cornell (1977), Mussa (1979) and Frenkel (1981b) have noted that exchange rate changes are largely unpredictable. Mussa (p. 10) states that: 'The natural logarithm of the spot exchange rate follows approximately a random walk.' The present study systematically confirms this 'stylized fact'. Another point Mussa makes and the results of this study support, is that any serial correlation found in the exchange rates by in-sample tests is likely to be unstable over time.

In our experiment, each competing model is used to generate forecasts at one to twelve month horizons for the dollar/pound, dollar/mark, dollar/yen and trade-weighted dollar exchange rates.^{2,3} The parameters of each model are estimated on the basis of the most up-to-date information available at the time of a given forecast. This is accomplished by using rolling regressions to re-estimate the parameters of each model every forecast period.

As representative structural models we choose the flexible-price monetary (Frenkel–Bilson) model, the sticky-price monetary (Dornbusch–Frankel) model, and the Hooper–Morton model. The latter empirical model extends the Dornbusch–Frankel model to incorporate the effects of the current account. We estimate these models using ordinary least squares, generalized least squares, and Fair's (1970) instrumental variables technique. We also try specifications which incorporate lagged adjustment.

A variety of univariate time series techniques are applied to the data as well. None of these techniques, including estimation of a random walk with drift model, generally yields any forecasting improvement in root mean square error or mean absolute error over the random walk model. Nor does an unconstrained vector autoregression composed of the exchange rate and all the explanatory variables from the structural models.

The forward rate does not predict any better than the random walk model either. But the interpretation of its relative performance is somewhat tangential to the main issue here, which is: How well do existing empirical exchange rate models fit out-of-sample?

A description of the competing models and the techniques used to estimate them is presented in section 2 of the paper. Section 3 discusses our methodology for comparing models out of sample, and section 4 contains the main results. In section 5 we list some possible explanations of these results.

2. A description of the models

Here we discuss the specification and statistical estimation of the various competing models.

2.1. *The structural models*

From the 'asset' models that have come to dominate the recent literature on exchange rate determination we select three which, perhaps due to the relative tractability of their data requirements, have been subjected to

²Haache and Townend (1981), using different methods than ours, conclude that none of the structural models can explain the behavior of the effective pound sterling exchange rate over the seventies.

³The trade-weighted dollar is a weighted-average of U.S. dollar exchange rates with the Group of Ten Countries plus Switzerland; see the data appendix.

extensive (in-sample) empirical testing. These are the flexible-price monetary (Frenkel–Bilson) model, the sticky-price monetary (Dornbusch–Frankel) model, and the sticky-price asset (Hooper–Morton) model.⁴ The quasi-reduced form specifications of all three models are subsumed in the general specification (1):

$$s = a_0 + a_1(m - \dot{m}) + a_2(y - \dot{y}) + a_3(r_s - \dot{r}_s) + a_4(\pi^e - \dot{\pi}^e) + a_5 \overline{TB} + a_6 \overline{\dot{TB}} + u, \quad (1)$$

where s is the logarithm of the dollar price of foreign currency, $m - \dot{m}$ the logarithm of the ratio of the U.S. money supply to the foreign money supply, $y - \dot{y}$ is the logarithm of the ratio of U.S. to foreign real income, $r_s - \dot{r}_s$ is the short-term interest rate differential and $\pi^e - \dot{\pi}^e$ is the expected long-run inflation differential.⁵ TB and $\overline{\dot{TB}}$ represent the cumulated U.S. and foreign trade balances, and u is a disturbance term. The disturbance term may be a serially correlated; we shall also consider allowing for lagged adjustment in eq. (1).

All of the models posit that, *ceteris paribus*, the exchange rate exhibits first-degree homogeneity in the relative money supplies, or $a_1 = 1$. The Frenkel–Bilson model, which assumes purchasing power parity, constraints $a_4 = a_5 = a_6 = 0$. The Dornbusch–Frankel model, which allows for slow domestic price adjustment and consequent deviations from purchasing power parity, sets $a_5 = a_6 = 0$. None of the coefficients in eq. (1) is constrained to be zero in the Hooper–Morton model. This model extends the Dornbusch–Frankel model to allow for changes in the long-run real exchange rate. These long-run real exchange rate changes are assumed to be correlated with unanticipated shocks to the trade balance.⁶ Imposing the constraint that domestic and foreign variables (except for trade balances) enter eq. (1) in differential form implicitly assumes that the parameters of the domestic and

⁴See Bilson (1978, 1979), Frenkel (1976), Dornbusch (1976), Frankel (1979, 1981), and Hooper and Morton (1982). Our nomenclature, which identifies particular models with authors who contributed significantly to their development is a conventional one. But it is not comprehensive in that some of these authors have worked with more than one of the three models, and there are other researchers who have studied these models or closely-related ones.

⁵Proxies for the unobservable $\pi^e - \dot{\pi}^e$ are typically constructed from variables such as long-term interest rate differentials, the preceding twelve-month period CPI or WPI inflation rates, or with an inflation rate autoregression; see Frankel (1981) and Hooper and Morton (1982).

⁶Since current account data is available only on a quarterly basis, the monthly version of Hooper and Morton's empirical model uses the trade balance as a proxy. Cumulative deviations from trend balances (current accounts) enter Hooper and Morton's equation since they assume that deviations from trend balances are unanticipated. Frankel (1982b) employs a model with a very similar quasi-reduced form, in which the cumulated current accounts of both countries enter because of wealth terms in the money demand equations. Branson, Halttunen and Masson (1979) also include the cumulated current accounts in empirical exchange rate equations. Their justification derives from the joint assumptions of imperfect asset substitutability and differential asset preferences across countries.

foreign money demand and price adjustment equations are equal. While this parsimonious assumption is conventional in empirical applications, it is a potential source of misspecification; see Haynes and Stone (1981). As reported below, however, no gain in out-of-sample fit results from estimating separate coefficients for domestic and foreign money supplies and real incomes.

A conventional approach to comparing the models subsumed in eq. (1) involves estimating a model of the general form, and then testing the constraints implied by the competing models. When overall performance is measured by in-sample fit, regressions based on eq. (1) do reasonably well. [See, for example, Frankel (1979) or Hooper and Morton (1982).] One drawback to this approach is that it is difficult to deal with the statistical problems encountered in obtaining consistent estimates of the coefficients in eq. (1). Variables such as relative money supplies and relative incomes are typically treated as exogenous variables in the underlying theoretical models, but may be more realistically thought of as endogenous variables. Other variables, such as the short-term interest differential, are generally endogenous in these same theoretical models. Yet they are still treated as legitimate regressors in ordinary or generalized least squares regressions of eq. (1).

The possibility that the explanatory variables in eq. (1) are endogenous is supported by the vector autoregression results presented in Meese and Rogoff (1983), and by the block exogeneity specification tests presented in Glaessner (1982). Of course, endogeneity of the explanatory variables does not preclude consistent estimation of the structural parameters in eq. (1). If, for example, the error term follows an autoregressive process of known maximum order, then instrumental variables techniques such as Fair's (1970) method are available. Frankel (1979) takes this approach, assuming a first-order autoregressive process and treating long-term expected inflation differentials as endogenous; Frankel (1981) allows short-term interest rates to be endogenous as well. To also account for the possible endogeneity of the money supplies, he tries constraining the coefficient on relative money supplies to its theoretical value of one. Following Frankel, we estimate the models using ordinary least squares, generalized least squares (correcting for serial correlation),⁷ and Fair's method. In the last case money supplies, short-term interest rates, and expected long-term inflation rates are treated as endogenous variables. Note that if one of the models summarized by eq. (1) is true, and if its structural and serial correlation parameters can be consistently estimated with instrumental variables techniques, then such techniques will outperform inconsistent techniques in large enough samples.

⁷Generalized least squares with a correction for a fifth-order autoregressive error term performs worse than GLS with a correction for a first-order autoregressive term (Cochrane-Orcutt). Cochrane-Orcutt also outperforms a stock-adjustment model.

However, generalized least squares parameter estimates did not yield inferior forecasts in our experiments.

While the forecasts generated in this study are based on models with freely-estimated coefficients, elsewhere [Meese and Rogoff (1983)] we try forecasting with the structural models using a grid of coefficient constraints drawn from the theoretical and empirical literature on money demand and purchasing power parity. Those results, which are discussed further below, do not lead to different conclusions than our experiments here with freely-estimated coefficients.

2.2. *Univariate and multivariate time series models*

Several univariate time series models involving a variety of prefiltering techniques and lag length selection criteria are employed in our experiments. All are estimated for the logarithm of the exchange rate.⁸

The prefiltering techniques involve differencing, deseasonalizing, and removing time trends. All six univariate time series techniques we consider are applied to both the actual and the prefiltered data. The first technique, the 'long AR', is an unconstrained autoregression (AR) where the longest lag considered (M) is a function of sample size (N), $M = N/\log N$. A deterministic rule like this has long been employed in spectral estimation [see Hannan (1970)] and has been applied to distributed lag models by Sims (1974b). If the true order of the autoregression is unknown but finite, this procedure is asymptotically inefficient relative to the Schwartz (1978) order selection criterion. We employ this procedure and the Akaike (1974) procedure in our study; the Schwartz criterion provides a consistent estimate of lag length, while the Akaike lag length criterion asymptotically produces minimum mean square prediction errors of the dependent variable. Our fourth procedure is like the long AR, except that in estimating the parameters more weight is given to recent observations. We arbitrarily choose to weight the observations by powers of 0.95. The fifth univariate technique involves direct application of the Wiener–Kolmogorov prediction formula in the frequency domain; see Sargent (1979).

A possible problem with all the techniques listed thus far is that they minimize criteria based on squared deviations. These type of criteria are inappropriate if, for example, exchange rates follow non-normal stable-Paretian distributions with infinite variance, as suggested by Westerfield (1977). Therefore, our final time series technique is based on minimizing absolute deviations. This 'MAD' estimator is more robust to fat-tailed distributions, and less sensitive to outlier observations.

⁸A Box–Cox transformation test indicates that the logarithmic transformation is slightly preferable to levels for all the exchange rates we consider. Theoretical reasons for preferring the logarithm of the exchange rate are given in section 3.

While the relative performance of the six univariate forecasting techniques is of interest in itself, we shall only report detailed results for the long AR model without trend, seasonal adjustment, or differencing. This model's performance characterizes those of the best univariate models; we will discuss the results of the other univariate models only to a lesser degree.

The random walk model, which uses the current spot rate as a predictor of all future spot rate is, of course, a univariate time series model. While the basic random walk model obviously requires no estimation, we also estimate a random walk model with drift parameter. The drift parameter is estimated as the mean monthly (logarithmic) exchange rate change.

An unconstrained vector autoregression (VAR), composed of the variables in eq. (1), serves as our representative multivariate time series model. A convenient normalization for estimation of the VAR is one in which the contemporaneous value of each variable is regressed against lagged values of itself and all the other variables, e.g. the exchange rate equation is

$$s_t = a_{i1}s_{t-1} + a_{i2}s_{t-2} + \cdots + a_{in}s_{t-n} + B'_{i1}X_{t-1} + B'_{i2}X_{t-2} + \cdots + B'_{in}X_{t-n} + u_{it}, \quad (2)$$

where X_{t-j} is a vector of the explanatory variables in eq. (1), lagged j periods. The error term u_{it} is serially uncorrelated, but may be contemporaneously correlated with the error terms in the other equations; thus the normalization used in (2) does not preclude contemporaneous interactions between variables. This normalization facilitates estimation since ordinary least squares equation by equation is an efficient strategy. The uniform lag length n across all (seven) equations is estimated using Parzen's (1975) lag length selection criterion.⁹ To reduce the parameterization of the VAR, we also try constraining the domestic and foreign cumulated trade balances to enter in differenced form rather than separately. As is not unusual in a small sample, the more parsimoniously parameterized six-variable VAR yields better forecasts, so this is the model we report below. The VAR is important to include in our forecasting experiments since it does not restrict any variables to be exogenous a priori, and is therefore robust to some of the estimation problems that plague the structural models discussed in the previous section. We did not, however, experiment with algorithms designed to reduce the number of estimated coefficients in the profligately parameterized VAR; see Litterman (1979). Such algorithms can sometimes markedly reduce the mean square prediction errors produced by these models.

⁹Parzen's (1975) criterion asymptotically selects a lag order greater than or equal to the true order, assuming the true order is finite. The lag lengths chosen for the dollar/mark, dollar/pound, dollar/yen, and trade-weighted dollar VARs are 2, 2, 4, and 2.

2.3. *Selecting the data*

The data, which are described in an appendix, are chosen to conform to the theoretical assumptions underlying the specification of the structural models.

All of the raw data used in this study are seasonally unadjusted, which makes it possible to estimate seasonal and structural parameters on a consistent basis. The use of seasonally adjusted data is especially likely to distort structural parameter estimates when the variables are not all adjusted by the same method. [See Sims (1974a, 1974b) for a further discussion.] We experimented with two different seasonal adjustment procedures. One method uses seasonal dummy variables. The other is Sims' (1974a) method which explicitly allows the seasonal parameterization to expand with sample size. As the results of our experiment are robust to the choice between these two techniques, we only report the results for the more conventional dummy variables procedure. Another reason for using seasonally unadjusted data is to avoid the use of certain information not available at the time of a given forecast. Forecasts based on seasonally adjusted data, adjusted over the extended sample period or with a two-sided filter such as Census X-11, implicitly make use of information which would not have been available.

The dollar/mark, dollar/pound, and dollar/yen spot exchange rate data are monthly point-sample data. We use an average of daily rates for the trade-weighted dollar, partly because that data is more readily available and partly to be consistent with other work on the trade-weighted dollar. [See Hooper and Morton (1982).] For the purposes of this study, point sample data have a decided advantage over monthly average data. Suppose the exchange rate follows a random walk on a mid-day to mid-day basis. Then as Working (1960) observed, a series consisting of monthly averages of mid-day rates will exhibit positive serial correlation.

Bilateral forward rates of one, three, six, and twelve month maturities are drawn from the same day of the month as the spot rates; point-sample short-term and (where possible) long-term interest rate data also match the spot rate data. We use treasury bill rates and interbank rates for short-term interest rates. Using these interest rates makes sense when estimating models based on standard money demand specifications. However, Euromarket rates would be more likely to conform to another assumption underlying most of the structural models: perfect asset substitutability. Some limited experimentation with Euromarket rates suggests that their use would have little effect on our results. As discussed below, the choice of monetary aggregates is potentially quite important. We try three different aggregates in our experiments: M1-B, M2, and the reserve-adjusted base. (Since the United Kingdom does not publish a series for M2, U.S. M3 and sterling M3 are employed in place of M2's for the dollar/pound rate experiments. Only M1-B

type measures are used in the trade-weighted dollar experiments; we were unable to find or construct a reserve-adjusted base series for Japan.)

3. The methodology for comparing models out-of-sample

All the competing models are estimated over a monthly data series which starts in March 1973, the beginning of the floating rate period, and extends through June 1981. Each model is initially estimated for each exchange rate using data up through the first forecasting period, November 1976. Forecasts are generated at horizons of one, three, six, and twelve months; these forecast horizons correspond to the available forward rate data. Then the data for December 1976 are added to the sample, and the parameters of each model, including the seasonal adjustment parameters, are re-estimated using rolling regressions. New forecasts are generated at one, three, six, and twelve month horizons, etc.

The purpose of considering multiple forecast horizons in this type of experiment is to see whether the structural models do better than time series models in the long run, when adjustment due to lags and/or a serially correlated error term has taken place. Of course, when lags and serial correlation are fully incorporated into the structural models, a consistently-estimated true structural model will outpredict a time series model at *all* horizons in a large sample.

The choice of where to begin forecasting is predicted on our desire to have sufficient degrees of freedom available for initial parameter estimates of all the models, especially the profligately-parameterized vector autoregression. We also look at the subperiod beginning in November 1978, in part because that date marks a major change in U.S. intervention strategy, and in part to see whether the relative performances of the competing models are different over the recent subperiod than over the entire forecasting period. Finally, we try truncating the sample in November 1980 which, like November 1976, marks a U.S. presidential election.

The structural models require forecasts of their explanatory variables in order to generate forecasts of the exchange rate. To give these models the benefit of the doubt, we use actual realized values of their respective explanatory variables. This procedure directly addresses one possible defense of these models: structural exchange rate models have explanatory power, but predict badly because their explanatory variables are themselves difficult to predict.¹⁰

The methodology used here for comparing models out-of-sample is drawn from the macro literature; see, for example, Nelson (1972), Christ (1975),

¹⁰When the explanatory variables are endogenous they will in general be correlated with the error term in eq. (1). If available, information about this correlation could be used to construct better structural model forecasts.

Litterman (1979), or Fair (1979). Although out-of-sample comparisons have considerable intuitive appeal, formal tests of whether these differences are statistically significant generally require restrictive assumptions.¹¹ But this limitation of the experimental design does not turn out to be crucial for the interpretation of our major result. We shall postpone this discussion until the next section.

Out-of-sample accuracy is measured by three statistics: mean error (ME), mean absolute error (MAE) and root mean square error (RMSE). These are defined as follows:

$$\text{mean error} = \sum_{s=0}^{N_k-1} [F(t+s+k) - A(t+s+k)]/N_k, \quad (3a)$$

$$\text{mean absolute error} = \sum_{s=0}^{N_k-1} |F(t+s+k) - A(t+s+k)|/N_k, \quad (3b)$$

$$\text{root mean square error} = \left\{ \sum_{s=0}^{N_k-1} [F(t+s+k) - A(t+s+k)]^2/N_k \right\}^{1/2}, \quad (3c)$$

where $k=1, 3, 6, 12$ denotes the forecast step, N_k the total number of forecasts in the projection period for which the actual value $A(t)$ is known, and $F(t)$ the forecast value. Forecasting begins in period t . Because we are looking at the logarithm of the exchange rate, these statistics are unit-free (they are approximately in percentage terms) and comparable across currencies. By comparing predictors on the basis of their ability to predict the logarithm of the exchange rate, we also avoid any problems arising from Jensen's inequality. Because of Jensen's inequality, the best predictor of the level of the dollar/mark rate might not be the best predictor of the mark/dollar rate.¹²

Root mean square error is our principal criterion for comparing forecasters. But because RMSE is an inappropriate criterion if, as mentioned

¹¹Granger and Newbold (1977, p. 281) propose a formal test of two forecasting techniques; the test is applicable only when both forecast errors are independent and normally distributed with zero means and constant variances. Thus, the test can only be applied at forecast intervals greater than one month if overlapping multi-horizon forecasts are omitted.

¹²Siegel (1972) notes that because $1/x$ is a convex function of the random variable x , $E(1/x)$ is not in general equal to $1/E(x)$. McCulloch (1975) suggests that this problem is not important empirically, given the historical variance of the exchange rate. Both analyses are based on an erroneous Taylor expansion which yields $E(1/x) - 1/E(x) \approx \text{var}(x)/E(x)$.³ This approximation may be misleading because the Taylor expansion used to derive it is local, whereas the expectations integral is global. While the above approximation is precisely correct when x follows a lognormal distribution, it can be way off when the distribution of x is skewed. Consider the discrete probability density function: $P(x=1)=0.99$, $P(x=0.01)=0.01$. Then $E(1/x) - 1/E(x) = 1.99 - 1.01 = 0.98$. However, $\text{var}(x)/E(x)^3 \approx 0.01$. The order of magnitude of the Jensen's inequality term is more likely to be large in data sets where an outside chance of a major intervention is incorporated into expectations.

in subsection 2.2, exchange rates are governed by a non-normal stable Paretian process with infinite variance, it is important to include mean absolute error. MAE is also a useful criterion when the exchange rate distribution has fat tails, even if the variance is finite. The last criterion, mean error, provides another measure of robustness. By comparing MAE and ME we can ascertain whether a model systematically over- or underpredicts.

4. The results

Table 1 lists the root mean square error statistics at one and twelve month horizons over the full November 1976 through June 1981 forecasting period for exchange rate for representative versions of each model. The structural models in table 1 are estimated using Fair's method as described in subsection 1.1. The measure of money is M1-B and, in the two sticky-price models, the long-term interest differential serves as the proxy for the long-run expected inflation differential. In addition, table 1 gives RMSE for the spot rate, the forward rate, the vector autoregression, and a long univariate autoregression (with order a function of sample size).

Ignoring for the present the fact that the spot rate does no worse than the forward rate, the striking feature of table 1 is that none of the models achieves lower, much less significantly lower, RMSE than the random walk model at any horizon. Although RMSE at three month horizons are not listed in table 1, they give the same result.

The structural models in particular fail to improve on the random walk model in spite of the fact their forecasts are based on realized values of the explanatory variables. They predict much worse, especially at one month horizons, if serial correlation is not accounted for. We obtain very similar results to those presented in table 1 using Granger and Newbold's (1977) method of combining the forecasts of structural models (without serial correlation) and time series models.¹³ Estimating the models in first difference form does not help, nor does using either a stock-adjustment formulation or generalized least squares with a correction for a fifth-order autoregressive process. Generalized least squares with a correction for first-order serial correlation (Cochrane–Orcutt) does frequently yield marginally better results than Fair's method. But this is not particularly encouraging

¹³Granger and Newbold's technique for optimally combining forecasts involves regressing the realized exchange rate against the forecasts of different models, with the weights constrained to sum to one, but not constrained to be positive. Even a bad predictor can sometimes be profitably combined with a good predictor; the forecasting gain depends on their covariation. An estimated combination of all seven forecasts never improves upon the random walk model alone, but estimated linear combinations of the different forecasts taken two at a time do sometimes outperform the random walk model. However, the same combination never works for more than one exchange rate. (These results are based on a November 1976–November 1980 forecasting period; linear combination forecasts were only generated at one month horizons.)

Table 1
Root mean square forecast errors.^a

Exchange rate	Model:	Horizon	Random	Forward	Univariate	Vector	Frenkel-	Dornbusch-	Hooper-
			walk	rate	autoregression	autoregression	Bilson ^b	Frankel ^b	Morton ^b
\$/mark		1 month	3.72	3.20	3.51	5.40	3.17	3.65	3.50
		6 months	8.71	9.03	12.40	11.83	9.64	12.03	9.95
		12 months	12.98	12.60	22.53	15.06	16.12	18.87	15.69
\$/yen		1 month	3.68	3.72	4.46	7.76	4.11	4.40	4.20
		6 months	11.58	11.93	22.04	18.90	13.38	13.94	11.94
		12 months	18.31	18.95	52.18	22.98	18.55	20.41	19.20
\$/pound		1 month	2.56	2.67	2.79	5.56	2.82	2.90	3.03
		6 months	6.45	7.23	7.27	12.97	8.90	8.88	9.08
		12 months	9.96	11.62	13.35	21.28	14.62	13.66	14.57
Trade-weighted dollar		1 month	1.99	N.A.	2.72	4.10	2.40	2.50	2.74
		6 months	6.09	N.A.	6.82	8.91	7.07	6.49	7.11
		12 months	8.65	14.24	11.14	10.96	11.40	9.80	10.35

^aApproximately in percentage terms.

^bThe three structural models are estimated using Fair's instrumental variable technique to correct for first-order serial correlation.

since Cochrane–Orcutt estimates are less often of the theoretically correct sign. Constraining the coefficients to be of the correct sign does not, however, improve the structural model forecasts. Meese and Rogoff (1983) report extensive constrained-coefficient experiments in which the structural models still fail to beat the random walk model at horizons of one to twelve months; see section 5 below.

Allowing for separate coefficients on domestic and foreign real incomes and money supplies yields no gain in out-of-sample forecasting accuracy. Nor does including domestic and foreign price levels as additional explanatory variables. Replacing M1-B with either M2 or the reserve-adjusted base almost always yields worse results, and never an improvement on the random walk model. Replacing long-term interest rates by other inflationary expectations proxies, such as current period inflation differentials, a moving average of past inflation differentials, or future inflation differentials, yields comparable results in both constrained- and freely-estimated coefficient experiments. The past twelve-month-period inflation differential works somewhat better (over the present sample). With that proxy, at one month horizons for the dollar/mark rate the Dornbusch–Frankel and Hooper–Morton models predict better than the random walk model in RMSE by 0.02 and 0.05 percent, respectively. They do worse at longer forecast horizons, though. The Hooper–Morton model (with the past year inflation rate proxy) also exhibits marginal improvement over the random walk model for the dollar/yen rate at six months (but not at one month or twelve months), and for the trade-weighted dollar at six months. At twelve months, the Hooper–Morton model improves by a more substantial 2 percent over the random walk model for the trade-weighted dollar.

The failure of the univariate time series models to beat the random walk model is similarly quite robust. None of the various univariate techniques improve on the random walk model at any horizon for the dollar/mark rate. The random walk with drift model improves by about 0.5 percent at six and twelve month horizons for the dollar/pound rate, but is 0.1 percent worse at one month and does much worse at predicting the dollar/yen and trade-weighted dollar. Given that we use monthly average data for the trade-weighted dollar, one would expect that in a large enough sample, the estimated univariate models would predict it better than the random walk model, even if the latter model is true for point-sample data (see subsection 2.3). But here only the long AR model (detrended) ever predicts the trade-weighted dollar better; it is only 0.5 percent better at twelve months. Only the Schwarz criterion together with detrending yields an improvement over the random walk model for the dollar/yen rate. The improvement is 1.5 percent at six months and 4.8 percent at twelve months.

It is worth emphasizing that even though exchange rates probably do not follow a random walk exactly, an estimated univariate model may not

forecast better due to sampling error. It is well known that imposing a coefficient restriction which is approximately correct tends to improve forecast accuracy. [See Sims (1980) or Litterman (1979) for further discussion.] This reason may similarly explain why the multivariate vector autoregression fails to outpredict the random walk model. It is thus possible that a non-structural method of reducing the number of estimated VAR parameters, such as the one Litterman (1979) proposes, would lead to an improvement on the random walk model.

At the risk of detracting from central issues such as how well existing empirical exchange rate models fit out-of-sample, we briefly turn to a comparison of spot and forward exchange rates. In table 1 the forward rate only improves on the random walk model in RMSE for the case of the dollar/mark rate at twelve month horizons. Given the joint assumptions of market efficiency and rational expectations, the relative performance of the forward rate may be interpreted as evidence on the existence of a risk premium.¹⁴ For example, the forward rate could predict worse than the random walk model when there is a time-varying risk premium, even if the risk premium is zero on average.

The dominance of the random walk model over the other models in RMSE remains when forecasting begins in November 1978, or alternatively if it ends in November 1980. The mean absolute error statistics, which are generally 20–25 percent smaller than RMSE, are not listed here since they yield virtually the same rankings as RMSE. Even the univariate technique designed to minimize mean absolute deviations fails to improve on the random walk model in out-of-sample MAE.

The mean forecast errors of the various models are listed in table 2. These errors are generally much smaller than the corresponding mean absolute errors, indicating that the models do not systematically over- or underpredict. (The structural models do tend to go systematically offtrack if no serial correlation is allowed for.) Note that the random walk model is somewhat less dominant in ME than in RMSE and MAE in our experiments, particularly for the dollar/mark rate. Estimating the structural models in first difference form generally produces lower ME but higher RMSE than does estimating the models in levels with a correction for serial correlation.

The results presented above do not answer the question of whether the random walk is *significantly* better than the other models in root mean square error, our primary criterion. However, given our finding that the

¹⁴A number of recent authors, including Bilson (1981), Cumby and Obstfeld (1981), Geweke and Feige (1979), Hakkio (1981), Hansen and Hodrick (1980, 1983), Meese and Singleton (1980), and Tryon (1979), have found evidence of the divergence of forward rates from expected future spot rates over the recent floating-rate period. Bilson (1981), however, is the only author who uses an out-of-sample testing methodology. Although his model is not discussed, it too failed to outperform the random walk model at one month forecast horizons.

Table 2
Mean forecast errors^a

Exchange rate	Model	Horizon	Random walk	Forward rate	Univariate autoregression	Vector autoregression	Frankel-Bilson ^b	Dornbusch-Frankel ^b	Hooper-Morton ^b
\$/mark		1 month	0.04	0.35	0.26	-1.12	0.37	-0.17	0.07
		6 months	-0.92	1.31	1.99	-3.31	1.23	-0.59	-0.17
		12 months	-3.93	0.29	5.20	-5.22	0.55	-3.06	-1.52
\$/yen		1 month	-0.46	-0.06	-0.15	-2.64	-1.36	-1.46	-0.18
		6 months	-3.32	-1.26	-3.17	-7.51	-8.00	-8.53	-1.81
		12 months	-6.48	-2.62	-8.91	-10.45	-14.05	-14.82	-2.38
\$/pound		1 month	-0.31	-0.38	-0.12	-3.72	-0.48	-0.37	-0.52
		6 months	-3.09	-4.05	-1.32	-9.45	-5.55	-4.53	-5.30
		12 months	-7.75	-9.55	-4.17	-18.54	-13.21	-12.07	-11.69
Trade-weighted dollar		1 month	-0.03	N.A.	0.06	0.89	0.63	0.54	0.68
		6 months	0.77	N.A.	1.61	3.91	3.86	2.79	3.52
		12 months	3.18	7.66	6.44	7.11	7.69	5.30	5.78

^aApproximately in percentage terms.

^bThe three structural models are estimated using Fair's instrumental variable technique to correct for first-order serial correlation.

random walk model almost invariably has the lowest root mean square error over all horizons and across all exchange rates, we *can* unambiguously assert that the other models do not perform significantly better than the random walk model. And while the random walk model may be as good a predictor as any of major-country exchange rates, it does not predict well. Even the RMSE in table 1 for the trade-weighted dollar — which as one might expect is more predictable than the bilateral rates — is 1.99 percent at one month and 8.65 percent at twelve months. The highest RMSE are for the dollar/yen rate: 3.70 percent at one month and 18.3 percent at twelve months.

One might hope to ultimately estimate a structural model which could perform substantially better than this, especially when forecasts are based on realized explanatory variable values. In the next section we address some possible explanations of our dissatisfying results. While the problem may lie in sampling error, it is also possible that these empirical models do not adequately capture expectations or other forces which influence exchange rates.

5. Possible reasons for the poor out-of-sample fit of the structural models

Since the structural model forecasts have been purged of explanatory variable uncertainty, their disappointing performance is most likely to be attributable to simultaneous equation bias, sampling error, stochastic movements in the true underlying parameters, or misspecification. Also, we make no attempt to account for possible non-linearities in the underlying models.

We have attempted to account for simultaneous equations bias by employing instrumental variables techniques, by estimating a vector autoregression, and by imposing theoretical coefficient constraints. The latter method is applied extensively in Meese and Rogoff (1983). There we develop and search a grid of coefficient constraints. The priors embodied in the grid are based on the fact that all the coefficients in the quasi-reduced form specification (1), except those on the cumulated trade balances, are functions of money demand parameters and the rate at which the real exchange rate returns to its long-run purchasing power parity level. [See Frankel (1979) or Hooper and Morton (1982).] Thus, we are able to base the coefficient on relative money supplies on the homogeneity postulate; the ranges for the income elasticity and interest rate semi-elasticity on the theoretical and empirical literature on money demand; and the range for the rate at which shocks to purchasing power parity are damped on empirical work on PPP.¹⁵ Despite allowance for a serially correlated error term we find that no element of the grid yields a constrained-coefficient forecaster which improves on the

¹⁵The money demand and PPP literature is discussed and cited in Meese and Rogoff (1983).

random walk model for horizons under twelve months. While there is sometimes sporadic improvement at longer horizons, the overall results are similar to those presented here. (We examine longer forecasting periods in our other study, since the constrained coefficient models require estimation of only the intercept term.) These further results appear to demonstrate that simultaneous equations bias and/or sampling error cannot be regarded as the primary rationalization of the evidence presented here.

Another candidate explanation for the poor performance of the structural models is that their underlying parameters shifted over the course of the seventies due to the effects of the two oil shocks, changes in global trade patterns, or changes in policy regimes. But unless the structural model parameters themselves follow a random walk, it does not necessarily follow that parameter instability can explain why the random walk model outperforms the structural models. Nevertheless, it may be fruitful to account for parameter instability by utilizing a method such as Kalman filtering, which weights recent observations more heavily in forming parameter estimates; see Sargent (1979). We did employ one univariate technique along these lines, weighted least squares, but as we reported above this isolated effort failed. It should be noted that there is a sense in which parameter instability is equivalent to having omitted (perhaps binary) variables.

A useful approach to investigating the problem of omitted variables, or misspecification in general, is to examine the (not strictly independent) building blocks of the structural exchange rate models subsumed in eq. (1): uncovered interest parity,¹⁶ the proxies for inflationary expectations, the goods market specifications, and the common money demand specification. Any or all of the above may be a source of misspecification; the discussion below is speculative.

Recent work on exchange rate risk premia has strongly challenged the assumption of uncovered interest parity. However, although the risk premia may be statistically significant, the evidence also suggests that the magnitudes are not large. (See the literature cited in footnote 14.) Therefore, it is not evident that deviations from uncovered interest parity can explain the poor forecasting performance of the structural models. [We note that empirical efforts to explain the risk premia in terms of portfolio-balance model variables have not been particularly successful; see, for example, Frankel (1982a).]

Measuring inflationary expectations presents many problems. The two sticky-price models, the Dornbusch–Frankel and Hooper–Morton models, are potentially quite sensitive to the proxy used for the long-run expected inflation differential. Proxies such as long-term interest rates and past inflation rates may be grossly inadequate. It is possible that the approach of

¹⁶Hooper and Morton allow for a risk premium as a function of central bank intervention in a more general version of their model.

estimating rational expectations versions of the models by imposing all the cross-equation restrictions, as in Driskell and Sheffrin (1981) or Glaessner (1982), will yield better expectations proxies. In that work, expectations of the exogenous forcing variables are formed using univariate or multivariate autoregressions. But it is not clear why autoregressions should necessarily yield good expectations proxies for the exogenous variables during a period when autoregressions yield poor proxies for the endogenous variables.

The goods market specifications of the models may also be suspect, though to differing degrees. There is little question that purchasing power parity did not hold in the short run during the seventies; see Isard (1977) or Frenkel (1981a). The Dornbusch–Frankel model assumes only long-run PPP; the evidence here is less clearcut [see Frenkel (1981b)]. The Hooper–Morton model attempts to empirically capture movements in the long-run real exchange rate, but it does not fit out-of-sample notably better than the other two models. Nevertheless, temporary or permanent movements in the PPP level of the exchange rate to real shocks may be a major cause of exchange rate volatility.

The final possible source of misspecification we shall discuss is the standard money demand function that underpins the models:

$$m - p = b_0 - b_1 r_s + b_2 y. \quad (4)$$

In (4), p is the logarithm of the price of the domestic good (using a different deflator for money balances would not alter the discussion below), and other variables are defined as in eq. (1). The breakdown of empirical money demand relationships is widespread, and the phenomenon is particularly acute for U.S. money demand equations; see Simpson and Porter (1980). In an attempt to control for unexplained shifts in velocity, we tried using eq. (4) and its foreign equivalent to substitute price levels for monetary variables in eq. (1). For the Frenkel–Bilson model the theoretical values of the coefficients in eq. (1) are the same as the corresponding coefficients in eq. (4), so price levels alone remain as regressors after the substitution. The transformed model is thus a purchasing power parity equation. In the two sticky-price models, the coefficient on short-term interest rates differs from the coefficient b_1 in (4). For those models, the coefficient on short-term interest rates in the quasi-reduced form exchange rate eq. (1) is the negative of the inverse of the goods market speed of adjustment parameter; see Frankel (1979). Therefore the price substitution only eliminates money supplies and real incomes in their quasi-reduced forms. In table 3 the models are estimated in the same fashion as in table 1, using Fair's method. The models still fail after the price levels substitution to improve on the random walk model in root mean square error. Implementing other estimation techniques and trying other expected inflation rate proxies yields qualitatively similar

Table 3
 Root mean square forecast errors with price levels substituted in for monetary variables.^{a, b}

Model:	Random walk	Modified Frenkel–Bilson (relative PPP)	Modified Dornbusch–Frankel	Modified Hooper–Morton	
Exchange rate	Horizon				
\$/mark	1 month	3.17	3.31	3.78	3.68
	6 months	8.71	9.78	11.28	10.51
	12 months	12.98	15.25	16.89	17.52
\$/yen	1 month	3.70	3.70	4.42	4.18
	6 months	11.58	12.55	13.86	12.69
	12 months	18.31	21.80	20.55	20.65
\$/pound	1 month	2.56	2.57	2.78	2.94
	6 months	6.45	7.83	8.64	8.92
	12 months	9.96	12.51	12.28	14.77
Trade-weighted dollar	1 month	1.99	2.18	2.29	2.48
	6 months	6.09	6.63	6.33	6.80
	12 months	8.65	10.48	9.52	9.45

^aApproximately in percentage terms.

^bThe three structural models are estimated using Fair's instrumental variable technique to correct for first-order serial correlation.

results to those reported in section 3, where the models are estimated without the price level substitution.

As a final effort to investigate the money demand problem, we estimated the Dornbusch–Frankel model for the three cross-exchange rates, thereby abstracting from the particularly unstable demand for money in the U.S. The Dornbusch–Frankel model does not outperform the random walk model for the pound/yen or pound/mark rates, but does do 0.6 percent better at six months (12.0 vs. 12.6 percent for the random walk) and 3.4 percent better at twelve months (16.0 vs. 19.4 percent) for the yen/mark cross-rate. But even this improvement is not so great as to provide a basis for asserting that money demand instability or misspecification is the main problem with the models.

6. Conclusions

The random walk model performs no worse than estimated univariate time series models, an unconstrained vector autoregression, or our candidate structural models in forecasting three major bilateral rates (the dollar/mark, dollar/pound, and dollar/yen) and the trade-weighted dollar. The results of

our paper contrast with those of previous studies based on in-sample fit. Thus, from a methodological standpoint, our paper supports the view that out-of-sample fit is an important criterion to consider when evaluating empirical exchange rate models.

The out-of-sample failure of the estimated univariate times series models and the vector autoregression suggests that major-country exchange rates are well-approximated by a random walk model (without drift). Of course, as long as the exchange rate does not exactly follow a random walk, we would expect one of the estimated time series models to prevail in a large enough sample.

Less certain is whether the failure of the structural models to outforecast the random walk model — even when uncertainty about the future values of the explanatory variables is removed — can similarly be attributed to sampling error. The constrained-coefficient experiments reported elsewhere in Meese and Rogoff (1983) suggest that neither sampling error nor simultaneous equations bias can fully explain the results presented here. We have listed other possible explanations without arriving at any definite conclusions. Structural instability due to the oil price shocks and changes in macroeconomic policy regimes, as well as the failure of the models to adequately incorporate other real disturbances, may be important. Misspecification of the money demand functions which underpin the structural models is another likely problem, although it is true that the structural models do not predict better when price levels are substituted in for monetary variables, or when M2 or the reserve-adjusted base are used in place of M1-B. Difficulties in modeling expectations of the explanatory variables are yet another obvious source of trouble. But determining the relative importance of the possible problems listed above, or any of the others listed in section 5, is at this point speculative.

Data appendix

The raw data consist entirely of seasonally unadjusted monthly observations over the period March 1973 to June 1981. In the bilateral data set for the United Kingdom, the spot and forward exchange rates, short-term interest rate, and long-term bond rate are all drawn from the same dates. Because daily long-term bond rate series are not readily available for Japan and Germany, only the exchange- and short-term interest rate dates correspond in those data sets. All other bilateral series as well as all of the series used in the trade-weighted data set are monthly data. All data are taken from publicly available sources. Data sources are listed below; a more detailed description of the data set can be found in Meese and Rogoff (1983).

United States data series

Long-term government bond yields, three-month Treasury bill rates, CPI,

industrial production, M1-B, M2 and the reserve-adjusted base: Federal Reserve Board data base.

Trade Balance: U.S. Department of Commerce, Highlights of U.S. Export and Import Trade.

Foreign data series for the bilateral data sets

Forward rates: Data Resources Inc. data base.

Spot rates: Federal Reserve Board data base.

Trade balances, monetary aggregates, and industrial productions: O.E.C.D. Main Economic Indicators.

Long-term bond yields and consumer prices: *Monthly Report of the Deutsche Bundesbank*, *Financial Times and Employment Gazette* (U.K.), *Economics Statistics Monthly* (Japan).

Interest rates (three-month): Frankfurter Allegemeine Zeitung (three-month German interbank rate), *Financial Times* (three-month British local authorities deposits), FRB data base ('Over two-month ends' bill discount rate, Tokyo stock exchange).

The trade-weighted data set

The weights utilized to determine the trade-weighted statistics are: German mark, 0.208; Japanese yen, 0.136; French franc, 0.131; United Kingdom pound, 0.119; Canadian dollar, 0.091; Italian lira, 0.090; Netherlands guilder, 0.083; Belgian franc, 0.064; Swedish Krona, 0.042; and Swiss franc, 0.036. These weights represent each country's share of the total trade (measured by the sum of imports plus exports) of all ten countries in the period 1972 through 1976. See Hooper and Morton (1978).

All the trade-weighted (foreign) data is drawn from O.E.C.D. Main Economics Indicators and the Federal Reserve data base.

References

- Akaike, H., 1974, A new look at the statistical model identification, *IEEE Transaction on Automatic Control* AC-19, 716-723.
- Amemiya, T., 1980, Selection of regressors, *International Economic Review* 21, 331-354.
- Bilson, John F.O., 1978, Rational expectations and the exchange rate, in: J. Frenkel and H. Johnson, eds., *The economics of exchange rates* (Addison-Wesley Press, Reading).
- Bilson, John F.O., 1979, The deutsche mark/dollar rate — A monetary analysis, in: Karl Brunner and Allan H. Meltzer, eds., *Policies for employment, prices and exchange rates*, Carnegie-Rochester Conference 11 (North-Holland Publishing Company, Amsterdam).
- Bilson, John F.O., 1981, The 'speculative efficiency' hypothesis, *Journal of Business* 54, 435-451.
- Branson, William H., Hannu Halttunen and Paul Masson, 1979, Exchange rates in the short run: Some further results, *European Economic Review* 12, 395-402.
- Christ, Carl F., 1975, Judging the performance of econometric models of the U.S. economy, *International Economic Review* 16, 54-74.
- Cornell, Bradford, 1977, Spot rates, forward rates and exchange market efficiency, *Journal of Financial Economics* 5, 55-65.

- Cumby, Robert and Maurice Obstfeld, 1981, A note on exchange-rate expectations and nominal interest differentials, *The Journal of Finance* 36, 697-703.
- Dornbusch, Rudiger, 1976, Expectations and exchange rate dynamics, *Journal of Political Economy* 84, 1161-1176.
- Driskell, Robert A. and Steven M. Sheffrin, 1981, On the mark: Comment, *American Economic Review* 71, 1068-1074.
- Fair, Ray C., 1970, The estimation of simultaneous equations models with lagged endogenous variables and first order serially correlated errors, *Econometrica* 38, 507-516.
- Fair, Ray C., 1979, An analysis of the accuracy of four macroeconomic models, *Journal of Political Economy* 87, 701-718.
- Frankel, Jeffrey A., 1979, On the mark: A theory of floating exchange rates based on real interest differentials, *American Economic Review* 69, 610-622.
- Frankel, Jeffrey A., 1981, On the mark: Reply, *American Economic Review* 71, 1075-1082.
- Frankel, Jeffrey A., 1982a, In search of the exchange rate risk premium: A six-currency test assuming mean variance optimization, University of California at Berkeley, mimeo.
- Frankel, Jeffrey A., 1982b, The mystery of the multiplying marks: A modification of the monetary model, *Review of Economics and Statistics* 64, 515-519.
- Frenkel, Jacob A., 1976, A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence, *Scandinavian Journal of Economics* 78, 200-224.
- Frenkel, Jacob A., 1981a, The collapse of purchasing power parities during the 1970's, *European Economic Review* 16, 145-165.
- Frenkel, Jacob A., 1981b, Flexible exchange rates, prices, and the role of news: Lessons from the 1970's, *Journal of Political Economy* 89, 665-705.
- Geweke, John and Edward Feige, 1979, Some joint tests of the efficiency of markets for forward foreign exchange, *Review of Economics and Statistics* 61, 334-341.
- Glaessner, Thomas, 1982, Theoretical and empirical essays on spot and forward exchange rate determination, Ph.D. Dissertation, University of Virginia.
- Granger, Clive and Paul Newbold, 1977, *Forecasting economic time series* (Academic Press, New York).
- Hacche, Graham and John Townend, 1981, Exchange rates and monetary policy: Modeling sterling's effective exchange rate, 1972-1980, in: W. Eltis and P. Sinclair, eds., *The money supply and the exchange rate* (Oxford University Press, Oxford).
- Hakkio, Craig S., 1981, Expectations and the forward exchange rate, *International Economic Review* 22, 663-678.
- Hannan, E.J., 1970, *Multiple time series* (John Wiley & Sons, New York).
- Hansen, Lars P. and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829-853.
- Hansen, Lars P. and Robert J. Hodrick, 1983, Risk averse speculation in forward exchange markets: An econometric analysis, forthcoming in: Jacob A. Frenkel, ed., *Exchange rates and international macroeconomics* (University of Chicago Press, Chicago).
- Haynes, Stephen E. and Joe A. Stone, 1981, On the mark: Comment, *American Economic Review* 71, 1060-1067.
- Hodrick, Robert J., 1979, On the monetary analysis of exchange rates: A comment, in: Kark Brunner and Allan H. Meltzer, eds., *Policies for employment, prices, and exchange rates*, Carnegie-Rochester Conference 11 (North-Holland, Amsterdam).
- Hooper, Peter and John E. Morton, 1978, Summary measures of the dollar's foreign exchange value, *Federal Reserve Bulletin* 64, 783-789.
- Hooper, Peter and John E. Morton, 1982, Fluctuations in the dollar: A model of nominal and real exchange rate determination, *Journal of International Money and Finance* 1, 39-56.
- Isard, Peter, 1977, How far can we push the law of one price?, *American Economic Review* 67, 942-948.
- Litterman, Robert B., 1979, Techniques of forecasting using vector autoregressions, Working paper no. 115, Federal Reserve Bank of Minneapolis.
- Maddala, G.S., 1977, *Econometrics* (McGraw-Hill, New York).
- McCulloch, J. Huston, 1975, Operational aspects of the Siegel paradox, *Quarterly Journal of Economics* 86, 303-309.

- Meese, Richard A. and Kenneth J. Singleton, 1980, Rational expectations, risk premia, and the market for spot and forward exchange, *International Finance Discussion Papers* no. 165, Board of Governors of the Federal Reserve System.
- Meese, Richard A. and Kenneth S. Rogoff, 1983, The out-of-sample failure of empirical exchange rate models: Sampling error or misspecification?, forthcoming in: Jacob A. Frenkel, ed., *Exchange rates and international macroeconomics* (University of Chicago Press, Chicago).
- Mussa, Michael, 1979, Empirical regularities in the behavior of exchange rates and theories of the foreign exchange market, in: Karl Brunner and Allan H. Meltzer, eds., *Policies for employment, prices and exchange rates*, Carnegie-Rochester Conference 11 (North-Holland, Amsterdam).
- Nelson, Charles, 1972, The prediction performance of the F.R.B.-M.I.T.-Penn. model of the U.S. economy, *American Economic Review* 62, 902-917.
- Parzen, Emanuel, 1975, Multiple time series: Determining the order of approximating autoregressive schemes, Technical Report no. 23, State University of New York at Buffalo.
- Sargent, Thomas J., 1979, *Macroeconomic theory* (Academic Press, New York).
- Schwarz, Gideon, 1978, Estimating the dimension of a model, *Annals of Statistics* 6, 461-464.
- Siegel, Jeremy J., 1972, Risk, interest rates, and the forward exchange, *Quarterly Journal of Economics* 86, 303-309.
- Simpson, Thomas D. and Richard D. Porter, 1980, Some issues involving the definition and interpretation of the monetary aggregates, in: *Controlling monetary aggregates*, vol. III, Federal Reserve Bank of Boston conference series no. 23.
- Sims, Christopher A., 1974a, Seasonality in regression, *Journal of the American Statistical Association* 69, 618-626.
- Sims, Christopher A., 1974b, Distributed lags, in: Michael D. Intriligator and David A. Kendrick, eds., *Frontiers of quantitative economics* (North-Holland, New York).
- Sims, Christopher A., 1980, Macroeconomics and reality, *Econometrica* 48, 1-48.
- Tryon, Ralph, 1979, Testing for rational expectations in foreign exchange markets, *International Finance Discussion Papers* no. 139, Board of Governors of the Federal Reserve System.
- Westerfield, Janice M., 1977, An examination of foreign exchange risk under fixed and floating rate regimes, *Journal of International Economics* 7, 181-200.
- Working, Holbrook, 1960, Note on the correlation of first differences in a random chain, *Econometrica* 28, 916-918.