Volume Title: Exchange Rates and International Macroeconomics

Volume Author/Editor: Jacob A. Frenkel, ed.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-26250-2

Volume URL: http://www.nber.org/books/fren83-1

Publication Date: 1983

Chapter Title: The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?

Chapter Author: Richard Meese, Kenneth Rogoff

Chapter URL: http://www.nber.org/chapters/c11377

Chapter pages in book: (p. 67 - 112)

3 The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?

Richard Meese and Kenneth Rogoff

3.1 Introduction

A companion study (Meese and Rogoff 1983) compared the out-of-sample fit of various structural and time-series exchange rate models and finds that the random walk model[1] performs as well as any estimated model at one- to twelve-month horizons for 1970s dollar/mark, dollar/pound, dollar/yen, and trade-weighted dollar exchange rates.[2] The structural models perform poorly even though their forecasts are purged of all uncertainty concerning the future paths of their explanatory variables by using actual realized values.

The present study demonstrates that the dismal short- to medium-run forecasting performance of the structural models is not attributable to the sample distribution of the coefficient estimates. We rule out that explana-

Richard Meese is a faculty member at the Berkeley School of Business. Kenneth Rogoff is an economist at the International Finance Division of the Board of Governors of the Federal Reserve System in Washington, D.C. Meese was also an economist at the Federal Reserve Board at the time this paper was written.

The authors have benefited from the comments and suggestions of Jeffrey Frankel, Robert Flood, Robert Hodrick, Peter Hooper, Peter Isard, and Julio Rotemberg. We are indebted to Julie Withers and Tamara McKann for excellent 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 or other members of its staff.

1. The structural models are described here in section 3.3 below. It should be noted that all the models considered are derivatives of the monetary or asset approach in that they specify real money demand at home and abroad as a function of real income, short-term interest rates, and possibly wealth. The random walk model predicts that today's exchange rate will obtain at all future dates.

2. In a study of the dollar/pound rate, Hacche and Townend (1981) use different methods to arrive at a similar conclusion; that the models do a very poor job of explaining the dollar/pound rate. The present study examines the three bilateral dollar rates and also cross-rates.

tion by showing that the models (with autoregressive error terms) perform poorly at one- to twelve-month forecast horizons over a wide range of coefficient values. These values are based on the theoretical and empirical literature on money demand and purchasing power parity. However, since the coefficient-constrained models only require estimation of the intercept terms, it is possible to look at longer forecast horizons here than in our other study. There the relative superiority of the random walk model over the structural models diminishes as the forecast horizon approaches twelve months. The present study explores the possibility that the structural models may improve on the random walk model forecasts at horizons of twelve to thirty-six months.

The main part of the paper is contained in section 3.3, which discusses the coefficient-constrained experiments. In section 3.2, vector autoregressions (VAR) are used to identify the factors that influence the exchange rate over short versus long horizons. The results from the VAR experiments also highlight the difficulties of finding legitimate instruments with which to estimate the structural models, thus motivating the constrained-coefficient approach of section 3.3. Section 3.4 asks which of the common building blocks of the structural models is most likely to have failed. We cite other studies and/or present evidence of our own on each of the factors which may have gone awry. But we are unable to clearly identify a single dominant cause of the breakdown of empirical exchange rate equations.

3.2 Decomposing the Forecast Error Variance of the Exchange Rate at Long and Short Horizons

Before proceeding to tests of the representative structural exchange rate models, we first examine a vector autoregression (VAR) consisting of the exchange rate and the explanatory variables of these models: relative money supplies, relative outputs,[3] relative short-term and long-term interest rates, and trade balances. The VAR is a tool for analyzing the relative importance of the explanatory variables in exchange rate model forecasts at both short and long horizons. As a by-product, the VAR also provides limited information on whether the conventional exogeneity assumptions used in estimation of the structural models are appropriate.

A convenient normalization for estimation of the VAR is one in which the contemporaneous value of each variable is regressed against lagged values of all the variables; for example, the exchange rate equation is given by

3. The assumption that U.S. and foreign variables enter exchange rate equation systems with equal but opposite signs is relaxed later in a limited number of experiments on the structural models. Economizing on variables in the otherwise highly parameterized VAR systems is quite important, so we only estimate the VAR models with the relative variables.

$$
(1) \qquad \begin{aligned} s_t = {} & a_{i1}s_{t-1} + a_{i2}s_{t-2} + \ldots a_{in}s_{t-n} \\ & + B'_{i1}X_{t-1} + B'_{i2}X_{t-2} \\ & + \ldots B'_{in}X_{t-n} + u_{it}, \end{aligned}
$$

where s_t is the (logarithm of the) exchange rate at time t and X_{t-j} is a vector of lagged values of the other included variables (listed above). Expressing the VAR system in the form of equation (1) facilitates estimation, as ordinary least squares equation by equation is an efficient estimation strategy. This normalization does not, however, preclude contemporaneous interactions between the variables, as these effects are captured in the covariance matrix of the disturbance terms u_{it}. The uniform lag length n across all (seven) equations is estimated using Parzen's (1975) lag length selection criterion.[4]

We estimate the VAR model for the dollar/mark, dollar/pound, and dollar/yen exchange rates over the floating rate period; the data consist of monthly observations for March 1973 through June 1981 (our seasonal adjustment procedures are described in appendix B). Once having obtained the coefficient estimates, the dynamic interactions among the variables are most easily studied with the use of the moving average (MA) representation, which is derived by inverting the autoregressive (AR) representation to express each of the endogenous variables in terms of the disturbances or innovations (the u_{it} in [1], for example). Studying the MA representation is complicated by the fact that the disturbance terms in the MA (or AR) representation are in general contemporaneously correlated (see Sims 1980 or Fischer 1982). So to simulate a "typical" shock to a given variable it is necessary to recognize that the expectation of other disturbances in the system, conditional on the particular shock of interest, is usually nonzero. Unfortunately, for two correlated disturbances $z_1(t)$ and $z_2(t)$, if $E[z_1(t)|z_2(t) = 1] = \alpha$, with α as an arbitrary constant, it is not in general true that $E[z_2(t)|z_1(t) = \alpha] = 1$. Because of this fact, there is no unique way to simulate "typical" shocks to these systems of endogenous variables when contemporaneous variable interactions are present. (In other words, when the covariance matrix of the disturbance terms is nondiagonal.) To identify a typical shock to the VAR system with a particular variable, we will follow the Sims (1980) procedure of specifying a variable ordering a priori. The variable ordering essentially specifies that the first variable is

4. Parzen's (1975) criterion selects an order ℓ^* which minimizes

$$
\mathrm{CAT}(\ell) = \mathrm{trace}\left(\frac{N}{T}\sum_{j=1}^{\ell} V_j^{-1} - V_\ell^{-1}\right), \ \ell = 1, 2, \ldots L,
$$

where N is the number of variables in the VAR, T is sample size, L is the maximal order considered, and V_j is an estimate (adjusted for degrees of freedom) of the covariance matrix of disturbances for the model with ℓ lags of each variable. Asymptotically, the order selected is never less than the true order, assuming the true order is finite.

predetermined with respect to all other variables, that the second variable is predetermined with respect to all but the first, etc. The identification of the VAR systems is pursued in greater detail in appendix A.

The multihorizon, forecast error variance decompositions listed in tables 3.1–3.3 are based on a variable ordering with the logarithm of U.S. to foreign relative money supplies, $m - m^*$, first, followed by the logarithm of relative outputs, $y - y^*$, the short-term interest differential, $r_s - r^*_s$, the long-term interest differential, $r_L - r^*_L$, the U.S. and foreign trade balances, TB and TB*, and the logarithm of the dollar price of foreign currency, s. In tables 3.4–3.6 the variable ordering is reversed. In the U.S.-German system, the largest estimated contemporaneous correlation is 44 percent between the short- and long-term interest differential equations. The other estimated contemporaneous correlations range from 5–20 percent. These results suggest that the variable ordering is potentially important in the U.S.-German VAR system. And indeed there are some differences between the U.S.-German VAR systems, regular versus reverse order, at both short- and long-forecast horizons. Note that in the regular (reverse) order system, exchange rate and long-term interest rate innovations account for 78.6 percent (93.7) and 12.8 percent (4.9) of the one-month-ahead forecast error variance of the exchange rate, and 48.1 percent (60.1) and 15.4 percent (10.5) of the thirty-six-month-ahead forecast error variance of the exchange rate. However, the statistical significance of these differences cannot be ascertained from tables 3.1 and 3.4. We have not yet performed the requisite (expensive) stochastic simulations to obtain estimates of the dispersion of the forecast error variance decompositions. Of course, the data necessarily contain less information about long-run variable interactions than short-run. Similar observations apply to the U.S.-U.K. and U.S.-Japanese VAR systems.

A second important observation to be made from tables 3.1–3.6 is that no variable appears to be exogenous to the VAR system. Abstracting from coefficient uncertainty, an exogenous variable would manifest itself as follows: at all horizons a variable's own innovations would account for all of its forecast error variance, so there would be a one in the column corresponding to a variable's own innovation and zeros elsewhere. (Block exogeneity is the obvious multivariate generalization.)[5] In the U.S.-German VAR, the exchange rate, relative incomes, the long-term interest differential, and the German and U.S. trade balances all appear to have large exogenous components, since for both variable orderings and all horizons (one to thirty-six months) own innovations in these variables explain at least 48, 55, 50, 59, and 65 percent of their respective

5. Glaessner (1982) formally tests the block exogeneity assumptions underlying some empirical specifications of the monetary model.

forecast error variances. For the U.S.-U.K. VAR, own innovations in the exchange rate, the U.K. and U.S. trade balances, and the long-term interest differential account for most of the forecast error variance of these variables. In the U.S.-Japanese system, it is the exchange rate, the Japanese and U.S. trade balances, and relative incomes that have this property.

The last feature of tables 3.1–3.6 that we wish to emphasize concerns the difference between those factors which appear to explain the forecast error variance of the three bilateral exchange rates at short horizons (one to three months) as opposed to longer horizons (one to three years). Based on the numbers reported in these tables it is clear that own innovations in exchange rates explain a large fraction of the exchange rate forecast error variance at one- and three-month forecast horizons, while innovations in the other variables become relatively more important at horizons of one and three years. This result is not atypical of VARs estimated on macroeconomic data (see Fischer 1982).

All of the features of tables 3.1–3.6 noted above suggest both (1) the difficulty in specifying the menu of variables to include in a structural exchange rate equation, and (2) the problems associated with finding legitimate instruments with which to consistently estimate the parameters of these models. The latter difficulty has led to the constrained-coefficient methodology of the next section.

3.3 Predicting and Explaining the Exchange Rate Out of Sample Using Structural Models with Constrained Coefficients

Elsewhere (Meese and Rogoff 1983) we employ rolling regressions to construct out-of-sample forecasts of the exchange rate using three structural models: a flexible-price monetary model (Frenkel-Bilson), a sticky-price monetary model (Dornbusch-Frankel), and a sticky-price asset model which incorporates the trade balance (Hooper-Morton).[6] The fact that these structural models do not outperform the random walk model at horizons of one to twelve months cannot be attributed to the inherent unpredictability of the explanatory variables; this uncertainty is purged from the forecasts by using realized explanatory variables values. Still, the possibility remains that our small-sample results can be attributed to poor parameter estimates rather than specification error. This

6. See Bilson (1978, 1979), Frenkel (1976), Dornbusch (1976*b*), Frankel (1979, 1981), and Hooper and Morton (1982). The identification of particular empirical models with authors who contributed significantly to their development follows one conventional nomenclature. It is relevant to note, however, that several of these same authors have analyzed more than one of the three models. For example, Frenkel (1981*b*) discusses a sticky-price model and emphasizes that the flexible-price model is a limiting approximation which is applicable in a highly inflationary environment. Dornbusch (1976*a*) examines a flexible-price monetary model with traded and nontraded goods.

Table 3.1 **Unconstrained U.S.-German VAR, March 1973–June 1981, Regular Variable Order (proportions of forecast error variance k months ahead attributable to each innovation[a])**

Forecast Error Variance in	k	Innovation in $m - m^*$	$y - y^*$	$r_s - r_s^*$	$r_L - r_L^*$	TB	TB*	s
$m - m^*$	1	.843	.005	.132	.004	.001	.004	.011
	3	.629	.003	.259	.029	.019	.040	.020
	12	.293	.007	.295	.100	.035	.249	.016
	36	.252	.007	.268	.122	.037	.254	.061
$y - y^*$	1	.006	.942	.000	.035	.011	.005	.001
	3	.006	.856	.002	.010	.008	.027	.001
	12	.014	.646	.004	.223	.029	.070	.014
	36	.015	.622	.004	.224	.031	.072	.032
$r_s - r_s^*$	1	.007	.024	.838	.073	.030	.007	.021
	3	.011	.017	.641	.080	.097	.124	.029
	12	.008	.016	.491	.081	.091	.282	.030
	36	.011	.017	.469	.088	.088	.276	.051

$r_L - r_L^*$	1	.024	.000	.161	.781	.002	.002	.030
	3	.056	.054	.126	.683	.007	.028	.044
	12	.152	.099	.101	.524	.036	.048	.040
	36	.147	.096	.108	.510	.039	.052	.048
TB	1	.080	.111	.009	.005	.769	.001	.025
	3	.078	.144	.011	.005	.721	.014	.026
	12	.068	.167	.010	.063	.617	.024	.050
	36	.068	.163	.010	.071	.596	.028	.063
TB^*	1	.021	.034	.063	.008	.042	.832	.001
	3	.024	.061	.047	.012	.036	.813	.006
	12	.023	.101	.037	.061	.032	.676	.069
	36	.028	.096	.040	.069	.030	.649	.086
s	1	.041	.011	.022	.128	.005	.006	.786
	3	.077	.008	.022	.163	.032	.010	.687
	12	.155	.034	.018	.125	.064	.050	.553
	36	.155	.033	.030	.154	.058	.090	.481

[a]Notes for tables 3.1–3.6: Columns of the table correspond to innovations in a particular variable for the specified forecast horizon $k = 1, 3, 12$, and 36. The rows add to one because the total forecast error variance attributable to each variable on the left of the table is allocated across the seven innovations. Abstracting from coefficient uncertainty, an exogenous variable would manifest itself as follows: At all horizons a variable's own innovations would account for all of its forecast error variance, so there would be a one in the column corresponding to a variable's own innovation and zeros elsewhere.

Table 3.2 **Unconstrained U.S.-Japan VAR, March 1973–June 1981, Regular Variable Order (proportions of forecast error variance k months ahead attributable to each innovation)**

Forecast Error Variance in	k	Innovation in $m-m^*$	$y-y^*$	$r_s-r_s^*$	$r_L-r_L^*$	TB	TB*	s
$m-m^*$	1	.943	.003	.001	.002	.012	.024	.014
	3	.866	.045	.004	.010	.024	.021	.029
	12	.551	.298	.021	.009	.045	.039	.038
	36	.424	.319	.070	.013	.048	.045	.080
$y-y^*$	1	.044	.929	.009	.001	.009	.000	.008
	3	.038	.915	.009	.008	.005	.020	.004
	12	.111	.637	.086	.019	.042	.090	.014
	36	.120	.531	.111	.019	.048	.072	.099
$r_s-r_s^*$	1	.142	.076	.770	.006	.002	.001	.003
	3	.183	.092	.581	.022	.030	.058	.034
	12	.110	.058	.284	.058	.069	.348	.073
	36	.115	.054	.238	.076	.120	.320	.076

$r_L - r_L^*$	1	.128	.014	.294	.520	.000	.003	.041
	3	.252	.039	.207	.342	.024	.022	.113
	12	.219	.046	.108	.184	.079	.092	.172
	36	.224	.071	.104	.172	.186	.085	.157
TB	1	.011	.016	.032	.000	.904	.005	.031
	3	.010	.022	.059	.011	.858	.008	.031
	12	.011	.032	.078	.017	.726	.097	.040
	36	.012	.036	.079	.024	.671	.111	.068
TB^*	1	.011	.036	.031	.064	.030	.823	.005
	3	.010	.045	.023	.095	.040	.780	.007
	12	.015	.051	.035	.105	.103	.626	.066
	36	.019	.041	.039	.093	.130	.523	.154
s	1	.011	.007	.007	.020	.071	.055	.828
	3	.016	.004	.004	.014	.139	.080	.741
	12	.016	.004	.003	.057	.285	.127	.507
	36	.026	.006	.003	.060	.313	.138	.455

Table 3.3 **Unconstrained U.S.-U.K. VAR, March 1973–June 1981, Regular Variable Order (proportions of forecast error variance k months ahead attributable to each innovation)**

Forecast Error Variance in	k	Innovation in $m-m^*$	$y-y^*$	$r_s-r_s^*$	$r_L-r_L^*$	TB	TB*	s
$m-m^*$	1	.911	.023	.059	.001	.006	.001	.000
	3	.789	.041	.142	.004	.004	.019	.001
	12	.398	.058	.292	.052	.004	.140	.057
	36	.270	.032	.200	.076	.028	.148	.245
$y-y^*$	1	.009	.820	.039	.012	.026	.027	.066
	3	.034	.600	.029	.087	.059	.073	.123
	12	.058	.431	.019	.138	.066	.183	.104
	36	.058	.407	.021	.145	.068	.177	.124
$r_s-r_s^*$	1	.047	.018	.880	.034	.001	.019	.000
	3	.034	.012	.830	.031	.001	.089	.003
	12	.136	.018	.641	.050	.000	.144	.009
	36	.159	.017	.600	.060	.002	.139	.024

$r_L - r_L^*$	1	.027	.007	.209	.725	.003	.018	.011
	3	.026	.032	.190	.637	.005	.066	.043
	12	.018	.091	.149	.528	.025	.050	.138
	36	.022	.086	.160	.496	.036	.051	.148
TB	1	.013	.041	.113	.016	.755	.018	.044
	3	.014	.050	.126	.069	.634	.057	.049
	12	.018	.052	.117	.078	.511	.082	.143
	36	.036	.047	.105	.074	.420	.115	.203
TB^*	1	.028	.069	.052	.002	.008	.830	.010
	3	.054	.106	.049	.038	.034	.710	.008
	12	.118	.128	.044	.068	.030	.596	.015
	36	.119	.116	.052	.079	.033	.555	.045
s	1	.000	.021	.014	.026	.007	.005	.926
	3	.001	.018	.008	.038	.008	.035	.891
	12	.012	.006	.002	.088	.048	.168	.675
	36	.066	.007	.058	.099	.059	.150	.560

Table 3.4 **Unconstrained U.S.-German VAR, March 1973–June 1981, Reverse Variable Order (proportions of forecast error variance *k* months ahead attributable to each innovation)**

Forecast Error Variance in		Innovation in						
	k	$m - m^*$	$y - y^*$	$r_s - r_s^*$	$r_L - r_L^*$	TB	TB*	s
$m - m^*$	1	.703	.017	.111	.047	.042	.017	.062
	3	.522	.011	.138	.191	.025	.058	.054
	12	.281	.024	.081	.298	.020	.264	.032
	36	.243	.022	.067	.287	.026	.264	.089
$y - y^*$	1	.011	.868	.007	.018	.068	.020	.009
	3	.009	.779	.014	.068	.057	.059	.013
	12	.014	.580	.049	.048	.075	.094	.096
	36	.015	.558	.050	.146	.076	.095	.059
$r_s - r_s^*$	1	.011	.016	.522	.350	.013	.085	.002
	3	.008	.013	.343	.333	.083	.115	.004
	12	.014	.018	.232	.279	.085	.353	.019
	36	.015	.018	.222	.272	.084	.343	.047

$r_L - r_L^*$	1	.000	.000	.002	.963	.008	.002	.024
	3	.013	.068	.013	.848	.014	.017	.026
	12	.080	.145	.014	.663	.048	.023	.027
	36	.081	.140	.015	.647	.051	.028	.038
TB	1	.011	.036	.000	.006	.918	.005	.023
	3	.014	.053	.008	.006	.873	.022	.023
	12	.017	.070	.024	.032	.757	.032	.068
	36	.018	.069	.025	.036	.733	.036	.083
TB^*	1	.001	.014	.001	.019	.028	.932	.004
	3	.003	.025	.009	.016	.023	.906	.017
	12	.003	.050	.034	.023	.024	.750	.117
	36	.008	.047	.035	.028	.024	.721	.137
s	1	.001	.005	.001	.049	.004	.003	.937
	3	.013	.007	.001	.068	.045	.007	.859
	12	.056	.013	.001	.060	.024	.054	.690
	36	.073	.013	.002	.105	.116	.091	.601

Table 3.5 **Unconstrained U.S.-Japan VAR, March 1973–June 1981, Reverse Variable Order (proportions of forecast error variance k months ahead attributable to each innovation)**

Forecast Error Variance in	k	Innovation in						
		$m - m^*$	$y - y^*$	$r_s - r_s^*$	$r_L - r_L^*$	TB	$\overset{*}{\mathrm{TB}}$	s
$m - m^*$	1	.749	.106	.028	.060	.013	.036	.010
	3	.686	.178	.020	.039	.034	.024	.019
	12	.394	.396	.023	.045	.064	.049	.024
	36	.271	.469	.045	.066	.051	.051	.045
$y - y^*$	1	.000	.898	.058	.002	.015	.019	.009
	3	.003	.857	.050	.016	.014	.044	.015
	12	.014	.693	.073	.088	.016	.092	.022
	36	.014	.658	.068	.090	.028	.071	.071
$r_s - r_s^*$	1	.011	.001	.505	.417	.016	.049	.001
	3	.056	.001	.503	.274	.022	.124	.020
	12	.046	.017	.308	.134	.052	.320	.124
	36	.052	.020	.273	.117	.107	.302	.130

$r_L - r_L^*$	1	.009	.003	.000	.826	.012	.122	.026
	3	.134	.002	.017	.599	.026	.118	.109
	12	.151	.029	.032	.304	.133	.105	.246
	36	.193	.068	.035	.283	.148	.097	.226
TB	1	.006	.016	.005	.007	.842	.009	.115
	3	.006	.017	.041	.006	.800	.010	.119
	12	.009	.017	.072	.010	.672	.114	.105
	36	.010	.029	.075	.012	.622	.131	.120
TB^*	1	.012	.008	.000	.001	.033	.928	.018
	3	.008	.013	.008	.003	.046	.906	.016
	12	.017	.035	.058	.006	.119	.739	.026
	36	.019	.038	.053	.012	.131	.601	.146
s	1	.001	.000	.002	.020	.000	.001	.976
	3	.004	.000	.004	.014	.024	.012	.942
	12	.009	.001	.024	.017	.137	.042	.739
	36	.016	.002	.027	.016	.175	.088	.676

Table 3.6 **Unconstrained U.S.-U.K. VAR, March 1973–June 1981, Reverse Variable Order (proportions of forecast error variance k months ahead attributable to each innovation)**

Forecast Error Variance in	k	Innovation in						
		$m - m^*$	$y - y^*$	$r_s - r_s^*$	$r_L - r_L^*$	TB	TB*	s
$m - m^*$	1	.749	.106	.028	.060	.013	.036	.010
	3	.686	.178	.020	.039	.034	.024	.019
	12	.394	.396	.023	.049	.064	.049	.024
	36	.271	.470	.045	.066	.051	.051	.045
$y - y^*$	1	.000	.898	.058	.002	.015	.019	.009
	3	.003	.857	.050	.016	.014	.044	.015
	12	.014	.693	.073	.088	.016	.092	.022
	36	.014	.658	.068	.090	.028	.071	.072
$r_s - r_s^*$	1	.011	.001	.505	.417	.016	.049	.001
	3	.056	.001	.503	.274	.022	.024	.020
	12	.046	.017	.308	.134	.052	.320	.124
	36	.052	.020	.273	.117	.107	.302	.130

$r_L - r_L^*$	1	.009	.003	.000	.826	.012	.122	.026
	3	.134	.002	.017	.594	.026	.118	.109
	12	.151	.029	.032	.304	.133	.105	.246
	36	.143	.068	.035	.283	.148	.097	.226
TB	1	.006	.016	.005	.007	.842	.009	.115
	3	.006	.017	.041	.006	.800	.010	.119
	12	.009	.014	.072	.010	.672	.114	.105
	36	.010	.029	.075	.012	.622	.131	.120
TB^*	1	.012	.008	.000	.001	.033	.928	.018
	3	.008	.013	.008	.003	.046	.906	.016
	12	.017	.035	.058	.006	.119	.739	.026
	36	.019	.038	.053	.012	.131	.601	.146
s	1	.001	.000	.002	.020	.000	.001	.976
	3	.004	.001	.004	.014	.024	.012	.942
	12	.009	.001	.024	.017	.137	.072	.739
	36	.016	.002	.027	.016	.175	.088	.676

possibility is especially worrisome in light of the estimated VAR models presented in the previous section. They indicate that it is difficult to find legitimate exogenous variables in the three structural exchange rate models. If this is the case, then consistent coefficient estimation becomes problematic and requires a priori knowledge of the serial correlation process of the error terms. These possible estimation problems may explain why the instrumental variables techniques implemented in our other study do not yield better results than ordinary least squares.

Here we explore a range of constrained coefficient models and present evidence that our previous results concerning one- to twelve-month forecast horizons cannot be explained by coefficient uncertainty. In addition, since the constrained coefficient models do not require a significant portion of the limited, floating rate data set for estimation, we are able to look at longer forecast horizons.

3.3.1 The Representative Structural Models

All three of the structural exchange rate models we consider are based on a common money demand specification, thereby allowing us to impose coefficient constraints on a consistent basis across models. The quasi-reduced form specification of each of the models is subsumed in the general specification below:

$$\begin{aligned} s = {} & a_0 + a_1(m - m^*) + a_2(y - y^*) + a_3(r_s - r_s^*) \\ & + a_4(\pi^e - \pi^{*e}) + a_5(\overline{\mathrm{TB}} - \overline{\mathrm{TB}}^*) + u, \end{aligned} \tag{2}$$

where $(\pi^e - \pi^{*e})$ is the expected long-term inflation differential, $\overline{\mathrm{TB}}$ and $\overline{\mathrm{TB}}^*$ are the cumulated U.S. and foreign trade balances, u is a disturbance term, and the other variables are as defined in section 3.2 above. (Recall that s is the logarithm of the dollar price of foreign currency.) In equation (2) we have imposed the usual constraint that domestic and foreign variables affect the exchange rate with coefficients of equal but opposite sign; this constraint is relaxed in a limited number of experiments both here and in our earlier study.[7] We choose not to specify an ad hoc lagged adjustment mechanism in (2), preferring to model the dynamics using an autoregressive error term as described below.

All three models hypothesize first-degree homogeneity of the exchange rate with respect to relative money supplies, or $a_1 = 1$. The Frenkel-Bilson or flexible-price monetary model, formed by differencing two identical money demand specifications while imposing purchasing

7. Haynes and Stone (1981) suggest that a problem with the representative structural models we consider is the restriction of equal but opposite coefficients on domestic and foreign variables. In Meese and Rogoff (1983), relaxing this restriction by letting certain domestic and foreign variables—incomes, money supplies, and cumulated trade balances—enter equation (2) separately yielded no forecasting improvement. Here we tried separating the incomes and trade balances, but again found no forecasting improvement.

power parity (PPP), posits the additional coefficient restrictions: $a_2 < 0$, $a_3 > 0$, and $a_4 = a_5 = 0$.

The Dornbusch-Frankel or sticky-price monetary model also hypothesizes that the coefficient on relative incomes $a_2 < 0$, but, in contrast to the Frenkel-Bilson model, hypothesizes that the coefficient on the short-term interest differential $a_3 < 0$, and that the coefficient on the long-term expected inflation differential $a_4 > 0$. The derivation of these coefficient restrictions is explained in Frankel (1979). The principal theoretical difference between the Frenkel-Bilson model and the Dornbusch-Frankel model is that the latter allows for short-run deviations from PPP caused by sticky domestic prices. Prices adjust only gradually in response both to excess demand, which depends on the terms of trade, and to secular inflation differentials ($\pi^e - \pi^{*e}$ in equation [2]). The long-run or flexible price exchange rate $\bar{s}$ is derived in the same manner as s in the Frenkel-Bilson model, except that it depends on $\pi^e - \pi^{*e}$, which is equal to the long-run, short-term interest differential:

$$\bar{s} = -\alpha + (m - m^*) - \varnothing(y - y^*) + \lambda(\pi^e - \pi^{*e}). \tag{3}$$

Using money demand functions of the form

$$m - p = \alpha - \lambda r_s + \varnothing y \tag{4a}$$

and

$$m^* - p^* = \alpha - \lambda r_s^* - \varnothing y^*, \tag{4b}$$

and a price adjustment equation of the form

$$(p - p^*)_{t+1} - (p - p^*)_t = \theta(s - p + p^*)_t + (\pi^e - \pi^{*e})_t, \tag{5}$$

Frankel demonstrates that augmented regressive expectations are rational:[8]

$$s_{t+1}^e - s_t = \theta(\bar{s} - s)_t + (\pi^e - \pi^{*e})_t, \tag{6}$$

where s_{t+1}^e is the exchange rate expected to prevail at time $t + 1$ based on period t information. Substituting (3) into (6) for $\bar{s}$, and also imposing uncovered interest parity by substituting $r_s - r_s^*$ for $s_{t+1}^e - s_t$, one arrives at the quasi-reduced form of the Dornbusch-Frankel model:

$$s = -\alpha + (m - m^*) - \varnothing(y - y^*) - \frac{1}{\theta}(r_s - r_s^*) + \left(\frac{1}{\theta} + \lambda\right)(\pi^e - \pi^{*e}).\text{[9]} \tag{7}$$

8. It is also straightforward to show that deviations from PPP caused by monetary shocks are expected to damp at rate θ.

9. Frankel uses both long-term interest differentials and past inflation differentials as proxies for $\pi^e - \pi^{*e}$, the flexible-price or long-run expected inflation differential.

So in the Dornbusch-Frankel model, a_3, the coefficient on the short-term interest differential $r_s - r_s^*$, does not depend on the nominal interest rate semielasticity of the demand for real balances λ. Rather it depends on the negative of the inverse of θ, the coefficient on excess demand in the price adjustment equation. The coefficient on the expected long-run inflation differential, a_4, is the sum of $1/\theta$ and λ.

The Hooper-Morton trade-weighted dollar model imposes the same constraints as the Dornbusch-Frankel model, except that it allows unanticipated shocks to the U.S. trade balance to affect the PPP or long-run real level of the exchange rate. In our bilateral version of their model, incipient trend U.S. trade balance surpluses require an appreciation of the long-run real exchange rate, while incipient trend foreign surpluses require a depreciation. Thus, $a_5 < 0$. It should be noted that the random walk model is also subsumed in the general specification (2). That model is given by $a_1 = a_2 = a_3 = a_4 = a_5 = 0$, and $u_t = u_{t-1} + e_t$, where e_t is a white noise process.

3.3.2 A Description of the Coefficient Constraints

The least controversial constraint we impose is that a_1, the coefficient on the logarithm of relative money supplies, is unity. While we shall not consider other values for a_1, we do experiment with different definitions of the money supply; the reserve adjusted base, M1-B, and M2 (in conjunction with their respective foreign counterparts).[10]

Widespread agreement is lacking on the values of the other parameters. For example, there is a range of theoretical and empirical estimates of the interest and income elasticities of money demand. The quantity theory puts the income elasticity at one, and the interest elasticity at zero. Alternatively, the Baumol (1952) and Tobin (1956) inventory theoretic approach, in its simplest form, can be used to derive an income elasticity of .5 and an interest elasticity of $-.5$. Taking into account integer constraints raises the income elasticity toward one and the interest elasticity toward zero (see Barro 1976). The Miller and Orr (1966) model of a firm's optimal cash-management procedures yields an interest elasticity of $-.33$. The income elasticity suggested by that model ranges from .33 to .67, depending on whether a rise in income brings a rise in the number of transactions or in the average size of transactions. The Whalen (1966) model of the precautionary demand for money also suggests an interest elasticity of $-.33$. In addition it yields an income elasticity which depends on how the size versus frequency of transactions changes as income rises, ranging from .33 to 1. Finally, we consider empirical estimates of the demand for money, for which Goldfeld's (1973) paper is a standard

10. For the dollar/pound rate we use M3's, since there are no data on M2 for the U.K. The results presented later in this section are based on M1 (M1 – B) data. However, we obtain very similar results with the different monetary aggregates.

reference. He estimates the income elasticity of money demand to be .19 in the short run and .68 in the long run; his short-run and long-run interest elasticities are −.064 and −.23. Since the present study takes the approach of modeling the serial correlation properties of the error term rather than specifying an ad hoc lagged adjustment mechanism, it follows that the higher long-run elasticities are more relevant for our purposes.

We are now ready to specify a complete grid of constraints for the Frenkel-Bilson model. The income elasticity constraints considered are .5, .65, .75, .85, and 1. This grid excludes the lowest ranges of income elasticities obtained in the theoretical models; we implicitly assume that income growth is accompanied by some growth in the size of transactions. The interest rate *semi*elasticity constraints include −3, −4.5, −6, −7.5, and −10. The latter grid encompasses interest rate elasticity priors ranging from somewhere between −.18 and −.21 to −.60 and −.70, depending on the bilateral exchange rate. The semielasticity priors are obtained by dividing the interest elasticity priors by the average prevailing level of short-term interest rates during the sample.

The grids of constraints for the Dornbusch-Frankel and Hooper-Morton models incorporate the same range of income elasticity and interest rate semielasticity constraints as the Frenkel-Bilson model grid. The two sticky-price models also require the specification of a grid for θ, the speed of adjustment parameter in the goods market. We choose a range of constraints for θ using the fact that it also represents the speed at which short-run deviations of the real exchange rate from its long-run equilibrium are damped. The grid for θ is based on the assumption that between 33 percent and 100 percent of today's deviation from PPP is expected to be eliminated one year hence. This range encompasses Genberg's (1978) estimates as well as those of Frankel (1979), both of which are based on data for Germany. Since in the Dornbusch-Frankel model the coefficient on relative short-term interest rates, a_3, is equal to $-1/\theta$, our grid of priors for a_3 in that model includes −1, −2, and −3. (The values for θ and therefore $-1/\theta$ are conceived on an annual basis, since the short-term interest differentials and expected long-term inflation differentials in the data set are annualized.) The coefficient on the expected long-run inflation differential, a_4, is equal to $\lambda + 1/\theta$, where $-\lambda$ is the interest semielasticity of money demand. The grid of constraints for a_4 is 4, 7, 9, 11, and 13, which includes the minimum and maximum possible values of $\lambda + 1/\theta$, given the individual grids of constraints for λ and $1/\theta$. For consistency, we exclude from our overall grid for the Dornbusch-Frankel model combinations of a_4 and a_3 such that $a_4 - a_3$ is less than 3 or greater than 10, the bounds on the grid for λ.

The coefficients on the cumulative monthly trade balances (taken as deviations from trend) in the Hooper-Morton model are based primarily on Hooper and Morton's work. We assume that a billion dollar U.S.

trade balance surplus above trend level leads, ceteris paribus, to an offsetting .3 to .5 percent appreciation of the dollar; that is, a_5, is .003 or .005. The results reported below are robust to using values of a_5 of .01 or .02. For simplicity, and to limit the size of the large grid of coefficient constraints for the Hooper-Morton model, we assume that a foreign trade balance surplus has an effect on the exchange rate of equal magnitude but opposite sign.

The final variable for which it is necessary to specify a grid of constraints is one which we logically know nothing about—the error term u_t. We assume that u_t follows a first-order autoregressive process:

$$u_t = \rho u_{t-1} + e_t = e_t/(1 - \rho L), \tag{8}$$

where e_t is white noise and L is the lag operator. The grid for the autoregressive parameter ρ is 0, .2, .4, .6, .8, and 1.0, so both the no serial correlation case and the first-difference case are covered.[11] The decision to analyze only a first-order autoregressive process is made in part to limit the size of the parameter grids, but it is also in part because of the results of our previous study. There, optimal linear combinations of structural models and very general autoregressive time series models are analyzed. A wide variety of optimal lag length selection criteria are used in developing the time series components of the forecasts; these criteria generally select a lag length of one for the univariate models.

Given the range of constraints we have selected, the grid for the Frenkel-Bilson model contains 150 different combinations of parameters; the grid for the Dornbusch-Frankel model has 330 elements; and the Hooper-Morton model grid has 660 elements.[12]

3.3.3 Results

The grid of parameter values developed above is now used to perform two basic experiments, designed to compare the structural models to the random walk model at forecast horizons of one, three, six, twelve, eighteen, twenty-four, thirty, and thirty-six months. The ex post and ex ante forecasting experiments differ mainly in whether forecasts are generated using realized values of the explanatory variables (ex post), or using predictions of the explanatory variables based on information available at the time of the forecast (ex ante).[13] The other difference is

11. For the Dornbusch-Frankel model we also experimented with a range of constraints on ρ concentrated between .8 and 1. The lowest end of this range produced the best results.

12. Recall that the Dornbusch-Frankel and Hooper-Morton model grids exclude combinations of a_3 and a_4 incompatible with the range of constraints specified for the interest rate semielasticity of real money demand λ.

13. In the absence of greater knowledge about the true underlying structure than is inherent in equation (2), it is not possible to take advantage of any correlation between the error term and the explanatory variables in generating the ex post forecasts. Such correlation is likely, though, given the endogeneity of the explanatory variables indicated by the VARs in section 3.2. In fact, if the variance of the error term is large and its (unknown)

that ex ante forecasting begins in June 1975 while ex post forecasting covers the entire sample period. The ex ante experiment requires enough observations for first-round estimation of the VAR that generates predictions of the explanatory variables. (Because the ex ante experiment is quite expensive to conduct, it is performed only for the Dornbusch-Frankel model.) Otherwise, the experiments are conducted in identical fashion. Constant terms corresponding to each constellation of parameter values are estimated using rolling regressions. The autoregressive component of forecasts made at time t are based on the period t error term.

The results of the ex post forecasting experiment are broadly characterized in table 3.7, where the structural model "forecasts" are compared with the random walk model forecasts on the basis of RMSE and MAE.[14] For each model and exchange rate, table 3.7 reports the shortest forecast horizon, in months, at which 0.1, 10, 25, and 50 percent of each model's parameter grid outpredicts the random walk model when realized values of the explanatory variables are used. Table 3.7 demonstrates that the results of Meese and Rogoff (1983) cannot be explained by parameter uncertainty. For the entire parameter grid and for all three exchange rates, the structural models never improve at all, much less significantly, on the random walk model in MAE or RMSE at forecast horizons less than twelve months. However, at horizons of twelve months or more—longer than we could examine in our study based on estimated coefficients—the RMSE and MAE of the models do sometimes improve on

covariance with the relevant linear combination of the explanatory variables (the "fundamentals") is negative, our ex post forecaster need not dominate optimal ex ante forecasters. In this perverse case, knowing that the realized fundamentals suggest a higher exchange rate means that you should guess a lower exchange rate.

14. The results for the Dornbusch-Frankel and Hooper-Morton models reported in tables 3.7–3.10 are obtained using long-term interest differentials as a proxy for expected long-run inflation differentials. It is important to recognize that these models are potentially quite sensitive to this variable. However, using instead current-period inflation differentials, a moving average of past inflation differentials, or future inflation differentials, yields qualitatively similar results in the ex post forecasting experiments (tables 3.7 and 3.8). We did not try these other proxies in the expensive ex ante experiments (tables 3.9 and 3.10).

Let $k = 1, \ldots, 36$ denote the forecast step, N_k the total number of forecasts in the projection period for which the actual value $A(t)$ is known, $F(t)$ the forecast value, and let forecasting begin in period $(t + 1)$. Define:

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

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

Our use of mean absolute error covers problems that might arise if, as suggested by Westerfield (1977), exchange rate changes are drawn from a stable Paretian distribution with infinite variance. The mean errors of the models (not reported) are small relative to mean absolute errors in almost all cases where $\rho > .2$, indicating that the structural models are not simply systematically over or underpredicting.

Table 3.7 Shortest Forecast Horizon (in months) for Which at Least x Percent of Each Model's Parameter Grid Improves on the Random Walk Model in MAE/RMSE When Realized Values of the Explanatory Variables Are Used[a]

Model	Threshold	Exchange rate:	dollar/mark		dollar/pound		dollar/yen	
		Metric:	MAE	RMSE	MAE	RMSE	MAE	RMSE
					(months ahead)			
	0–1%		24	30	18	24	12	12
Frenkel-Bilson	10		30	30	18	24	18	18
(grid size = 150)	25		30	30	24	30	24	24
	50		36	36	30	36	36	30
	0–1%		12	18	18	18	12	12
Dornbusch-Frankel	10		18	18	24	24	12	12
(grid size = 330)	25		30	30	30	36	12	12
	50		—	—	—	—	24	18
	0–1%		12	18	18	18	12	12
Hooper-Morton	10		18	18	24	24	12	12
(grid size = 660)	25		30	30	30	36	12	18
	50		—	—	—	—	24	18

[a]MAE is mean absolute error; RMSE is root-mean-square error. The text contains a description of the parameter grids. To read table 3.7, consider the entry in the second row, first column (30). The earliest forecast horizon at which at least 10 percent of the Frenkel-Bilson model parameter grid can improve on the random walk model in MAE for the dollar/mark rate is thirty months out.

the random walk model. This result is tempered by the fact that the minimum RMSE or MAE coefficient configurations bounce around at different forecasting horizons. Still, the percentage of the parameter grids which improve on the random walk model does increase with forecast horizon. Overall these essentially in-sample results—in-sample because not all coefficient configurations improve on the random walk model—must be interpreted with caution.

Table 3.8 presents best representative parameter values for each of the models, together with their corresponding RMSE and MAE.[15] These two statistics are also given for the random walk model. At thirty-six months, the best representative coefficient values for the Dornbusch-Frankel and Hooper-Morton models do about 50 percent better than the random walk model in RMSE and MAE; the Frenkel-Bilson model only does about 30 percent better.

Since the models do not forecast well at short horizons in the ex post experiment, it is not surprising that the one model considered in the ex ante experiment does poorly at short horizons as well.[16] Tables 3.9 and 3.10 present results for the ex ante forecasting experiment with the Dornbusch-Frankel model. No parameterization of that model ever improves on the random walk model in MAE for horizons under twelve months; the threshold horizon is even longer when RMSE is the metric. Furthermore, for the dollar/pound and dollar/yen exchange rates, over 90 percent of the parameter grids fail to beat the random walk model in MAE or RMSE at any horizon. It is true, however, that at thirty-six months the best representative Dornbusch-Frankel model performs almost as well in the ex ante experiment as the best representative Dornbusch-Frankel model in the ex post experiment (compare tables 3.8 and 3.10). Again, we should emphasize that the evidence presented here on the possible forecasting superiority of the structural models is essentially in-sample, since not all configurations of the parameter constraints improve on the ramdom walk model.

Also reported in table 3.10 are the forecasting properties of the seven-variable VAR system of section 3.2. This model, estimated by rolling regressions, is a true ex ante forecaster. The VAR outforecasts the random walk model at three-year horizons for the dollar/mark rate. It does worse at one-year horizons for that exchange rate, though, and

15. The "best" representative set of parameter constraints for each model in table 3.8 is chosen in an ad hoc fashion as the one which comes in first (also ahead of the random walk model) at the greatest number of horizons. The maximum improvements over the random walk model in MAE or RMSE at thirty-six-month horizons exhibited by these representative models are as large as those exhibited by any other parameter configurations.

16. While only one model is considered in the ex ante experiment, note that all three models yield qualitatively equivalent results for the ex post experiment. Also, since the Dornbusch-Frankel model predicts that the exchange rate will return in the long run to its flexible-price or Frenkel-Bilson model value, we should a priori expect the performance of both models at long forecast horizons to be quite similar.

Table 3.8 **Comparing the Random Walk Model and the Structural Models (with their best representative parameter configurations) When Realized Values of the Explanatory Variables Are Used**[a]

Model	Horizon	dollar/mark		dollar/pound		dollar/yen	
		MAE	RMSE	MAE	RMSE	MAE	RMSE
Random walk	1	2.4	3.2	2.0	2.5	2.1	3.0
	3	4.8	6.2	3.2	5.1	4.2	5.7
	12	9.4	10.9	9.8	11.5	10.6	13.8
	36	18.1	21.0	23.4	25.4	19.4	23.3
(a_2, a_3, ρ)		$(-.5, 4.5, .4)$		$(-1, 3, .8)$		$(-.5, 4.5, .8)$	
Frenkel-Bilson	1	9.1	11.4	4.2	6.1	4.5	6.4
	3	11.5	14.2	8.7	11.1	7.9	11.5
	12	12.2	15.2	13.5	16.6	9.7	13.3
	36	12.6	17.0	15.5	18.8	10.2	14.5
(a_2, a_3, a_4, ρ)		$(-.85, -1, 6, .4)$		$(-.5, -1, 4, 0)$		$(-.5, -1, 9, .8)$	
Dornbusch-Frankel	1	5.5	6.9	8.1	10.0	4.4	8.4
	3	8.1	9.7	8.6	10.5	7.4	9.4
	12	8.8	10.8	10.4	12.3	7.0	8.5
	36	8.2	10.5	8.3	10.0	8.8	10.2
$(a_2, a_3, a_4, a_5, \rho)$		$(-.5, -1, 7, .005, 0)$		$(-.5, -1, 4, .005, 0)$		$(-1, -1, 9, .003, .8)$	
Hooper-Morton	1	8.3	10.4	4.0	10.0	4.9	6.7
	3	8.8	11.0	8.5	10.5	8.3	10.9
	12	9.2	11.6	10.0	12.0	8.8	11.7
	36	9.3	11.6	10.5	12.5	9.3	12.2

[a]MAE (mean absolute error) and RMSE (root-mean-squared error) are approximately in percentage terms, since "forecasts" are for the logarithm of the exchange rate. Forecasts are compared over the period March 1973–June 1981.

Table 3.9 **Shortest Forecast Horizon (in months) for Which at Least x Percent of the Dornbusch-Frankel Model's Parameter Grid Improves on the Random Walk Model in MAE/RMSE When Predicted Values of the Explanatory Variables Are Used[a]**

		dollar/mark		dollar/pound		dollar/yen	
Model	Threshold	MAE	RMSE	MAE	RMSE	MAE	RMSE
				(months ahead)			
	0–1%	12	24	30	36	12	18
Dornbusch-Frankel	10	12	24	—	—	—	—
(grid size = 330)	25	36	36	—	—	—	—
	50	—	—	—	—	—	—

[a]The description of table 3.9 is essentially the same as given for table 3.7. However, in contrast to the experiment of table 3.7, in which realized values of the explanatory variables are used to generate forecasts of the exchange rate, table 3.9 is based on an experiment in which the explanatory variables are forecast with a VAR.

Table 3.10 **Comparing the Random Walk Model, a VAR Model Estimated by Rolling Regressions, and the Best Representative Dornbusch-Frankel Model When Predicted Values of the Explanatory Variables Are Used**[a]

Model	Horizon	dollar/mark		dollar/pound		dollar/yen	
		MAE	RMSE	MAE	RMSE	MAE	RMSE
Random walk	1	2.2	3.0	2.0	2.6	2.3	3.2
	3	4.0	5.2	4.1	5.2	4.7	6.2
	12	10.1	11.7	10.1	11.5	13.2	16.1
	36	24.2	26.2	18.8	21.2	24.8	28.1
(a_2, a_3, a_4, ρ)		$(-.5, -1, 4, 1.0)$		$(-.5, -1, 4, 0)$		$(-.5, -3, 11, .8)$	
Dornbusch-Frankel	1	2.4	3.2	14.0	17.3	2.9	4.2
	3	4.6	5.7	14.7	18.0	6.6	8.6
	12	7.3	10.9	17.5	20.1	12.4	16.0
	36	12.4	19.1	9.0	10.8	17.0	18.7
Unconstrained VAR	1	5.2	6.3	5.4	6.3	4.9	6.4
	3	7.9	9.5	8.0	9.6	7.4	9.5
	12	11.1	13.2	17.3	19.3	15.9	19.6
	36	16.9	18.5	39.5	44.8	37.5	40.6

[a]MAE (mean absolute error) and RMSE (root-mean-squared-error) are approximately in percentage terms, since forecasts are for the logarithm of the exchange rate. Forecasts are compared over the period June 1975–June 1981.

worse at all horizons for the dollar/pound and dollar/yen exchange rates. It is possible that these results can be improved by imposing probabilistic priors on the VAR (see Litterman 1979). (An identified structural model such as the Dornbusch-Frankel model can be thought of as a VAR with a priori restrictions.)

3.4 The Poor Performances of the Structural Models: Possible Causes

The constrained coefficient experiments of section 3.3 reinforce the results of our earlier study. The selected structural models (with autoregressive error terms) fail to forecast or even explain out of sample as well as the random walk model at horizons of up to twelve months. The models do sometimes produce better forecasts than the random walk model at longer horizons, but in an unstable fashion. As noted in section 3.2, the limited floating rate data set necessarily contains more information about short than about long forecast horizons.

In this section we try to trace the instability or misspecification of these empirical exchange rate equations to their building blocks, such as uncovered interest parity,[17] the particular money demand specification, the proxies for inflationary expectations, and the goods market specifications. These building blocks are not, of course, strictly independent.

The assumption of uncovered parity has been strongly challenged by recent work on exchange rate risk premiums.[18] However, while some authors find evidence of risk premiums, the weight of the evidence is that the magnitudes involved are not large. Nevertheless, volatile time-varying risk premiums remain a possible explanation of the results.

The goods market specifications of the three representative structural models are relatively simple. The Frenkel-Bilson flexible-price monetary model imposes PPP, even in the short run. The Dornbusch-Frankel sticky-price monetary model allows for short-run deviations from PPP. The Hooper-Morton model is similar except that it attempts to incorporate movements in the long-run PPP level of the exchange rate by assuming that these movements take place in response to unanticipated trade balance (current account) deficits or surpluses. While short-run PPP does not provide an accurate characterization of the 1970s,[19] there is

17. In some versions of the Hooper-Morton model this assumption is relaxed.

18. See, for example, Hansen and Hodrick (1980*a*, *b*), Cumby and Obstfeld (1981), Hakkio (1981), Tryon (1979), Bilson (1981), Meese and Singleton (1982), or Geweke and Feige (1979). Hansen and Hodrick study this issue in their paper contained in this volume.

19. Isard (1977), Genberg (1978), and Frenkel (1981*a*, *b*) provide evidence on this point. In this context, it would be useful to remind the reader that our identification of particular models with particular authors oversimplifies the history, development, and application of these models (see note 6).

no strong evidence that the long-run PPP level of the exchange rate changed significantly.

The performance of the Dornbusch-Frankel and Hooper-Morton models are potentially quite sensitive to the use of a variable other than the long-term interest differential as a proxy for the long-run expected inflation differential. Although we did not find a proxy which yielded better results (see note 14), this issue merits further attention.

Perhaps the major problem with the structural models considered here is the instability of their underlying money demand specifications. The recent breakdown of U.S. money demand relationships was first noted by Goldfeld (1976) and is documented extensively by Simpson and Porter (1980). Conventional empirical money demand specifications such as equations (4) of section 3.3 have consistently underpredicted U.S. M1 velocity since mid-1974. For this reason, the present study uses M1-B, for which the systematic bias over the sample period is much smaller, and the new definition of M2, for which the bias is negligible. But equations (4) still fail to predict these aggregates or the reserve-adjusted base with any notable degree of precision. As reported above, our exchange rate experiment results are not sensitive to which of these monetary aggregates (together with their respective foreign counterparts) we employ.

Whether or not money demand instability and/or misspecification is responsible for the exchange rate results, it is certainly true that the conventional money demand equation does not work well when expressed in terms of U.S. minus foreign variables. Equation (4a) minus (4b) fails Chow (1960) tests for the stability of the intercept term at four different breaks in the sample. It also fails Goldfeld and Quandt (1965) tests of homoscedastic disturbance terms over the sample breaks.[20]

To investigate the possibility that our results are generated solely by money demand instability in the United States, we performed ex post forecasting experiments using the Dornbusch-Frankel model on the pound/mark, pound/yen, and yen/mark cross-exchange rates. For the case of the yen/mark, we found coefficient values for which the model pulled even with the random walk model as early as six months. But the subsequent improvement at longer horizons never exceeded 30 percent. (The pound/mark and pound/yen cross-rate results are no better than the results for the various dollar exchange rates.)

In sum, money demand instability is an important potential explanation for our results, but further work is needed to demonstrate that time-varying risk premiums, volatile long-run real exchange rates, or

20. The breaks in the sample at which these stability tests are conducted are chosen arbitrarily and correspond to (1) June 1974—the start of the mature float; (2) November 1976—the approximate sample midpoint; (3) November 1978—the dollar support program; and (4) October 1979—the change in Federal Reserve operating procedures. The tests were conducted using all parameter configurations with the grids for (λ, $\emptyset$) reported in section 3.3.2.

poor measurement of inflationary expectations are not the dominant problems.

3.5 Conclusions

The unimpressive out-of-sample performance of the Frenkel-Bilson, Dornbusch-Frankel, and Hooper-Morton empirical exchange rate models cannot be attributed to inconsistent or inefficient parameter estimates. These models fail to yield any improvement over the random walk model in mean absolute or root-mean-squared error one to twelve months out of sample for a broad range of theoretically plausible coefficient values, even when autoregressive error terms are introduced. Thus it is unlikely that more efficient estimation techniques, such as imposing all the cross-equation rational expectations restrictions, will yield parameter estimates which do better.[21] The constrained coefficient models do prevail at longer horizons but in an unstable fashion; the best coefficient values bounce around depending on the forecast horizon.[22]

The breakdown of empirical exchange rate models may be the result of volatile time-varying risk premiums, volatile long-run real exchange rates, or poor measurement of inflationary expectations. Alternatively, the main problem may lie in their money demand specifications. At this point, we are reluctant to draw any firm conclusions.

Appendix A

In this appendix we describe the triangularization of the VAR system used in section 3.2 to analyze the dynamic effects of an innovation to a particular variable. First suppose the tth observation of the VAR is represented by

$$\text{(A1)} \qquad [I_N - A(L)]y_t = u_t,$$

where $[I_N - A(L)]$ is a matrix polynomial in the lag operator L, y_t is the $N \times 1$ vector of variables in the system, $E(u_t) = \mathbf{0}$, and $\text{var}(u_t) = V$, positive definite. Using the Cholesky factorization $V = WW'$, where W is lower triangular, we can transform (A1) to the system

$$\text{(A2)} \qquad W^{-1}[I_N - A(L)]y_t = W^{-1}u_t = e_t,$$

21. See Driskill and Sheffrin (1981) or Glaessner (1982). These more sophisticated statistical techniques may provide superior expectations proxies, however.

22. If the true structural model were known and combined with an accurate representation of the serial correlation process of the error term, then such a model would produce minimum MSE forecasts at all horizons.

where $E(e_t) = \mathbf{0}$ and $\text{var}(e_t) = I_N$, the order N identity matrix. Since W^{-1} is lower triangular, the system (A2) is recursive, as described in the text. The moving average representation of (A2) is

$$\text{(A3)} \qquad y_t = [I_N - A(L)]^{-1} W e_t,$$

and in this expression the contemporaneous value of the first component of e enters all N equations, the contemporaneous value of the second component of e enters the last $N - 1$ equations, etc. Because the decomposition of V is not unique, studying the effect of the uncorrelated innovations e_t on y_t will depend on the variable ordering unless V is diagonal, that is, unless the system (A2) has no contemporaneous interactions among variables.

Expression (A3) is also used to construct the variance decompositions of tables 3.1–3.6. Since all components of e_t have unit variance, the variance of y_{it} (the ith element of the vector y_t) is the sum of squares of the elements in the ith row of $[I_N - A(L)]^{-1}W$. The percentage of the forecast error variance of y_{it} explained by the jth innovation e_{jt} (the jth element in the vector e_t) is calculated as the ratio of the sum of squares of the (i, j) element of $[I_N - A(L)]^{-1}W$ to the variance of y_{it}.

Appendix B Data Sources

The data set consists of seasonally unadjusted monthly observations over the period of March 1973 to June 1981. All the raw data are seasonally adjusted using dummy variables (the results reported in the text are insensitive to the use of more sophisticated seasonal adjustment procedures described in Meese and Rogoff 1983). In the U.K. data set, the spot and forward exchange rates, short-term interest rate, and long-term bond rate are always drawn from the same date. Because daily bond series are not readily available for Japan and Germany, only the exchange and interest rates correspond in these data sets. All other series are monthly data, and all data are taken from publicly available sources.

The bilateral data sets draw exchange rate data from identical sources, as follows:

One-, Six-, and Twelve-Month Forward Exchange Rates

Data Source: Data Resources, Inc. data base.

Series: One-, six-, and twelve-month forward bid rates in U.S. dollars per local currency unit.

Description: Daily data based on 10:00 A.M. opening New York market rates.

Three-Month Forward and Spot Exchange Rates

Data Source: Federal Reserve Board data base.

Series: Three-month forward and spot bid rates in U.S. dollars per local currency unit.

Description: Daily data based on 12:00 noon New York market rates.

Sources of the other data series are discussed below by country.

Germany

Bond Yields

Data Source: Deutsche Bundesbank, *Statistical Supplement to the Monthly Reports of the Deutsche Bundesbank*, series 2, Securities Statistics, table 7b.

Series: Yields in percent per annum on fully taxed outstanding bonds of the Federal Republic of Germany.

Description: Monthly data. Data are calculated as averages of four bank-week return dates including the end-of-month yield of the preceding month.

Consumer Prices

Data Source: Deutsche Bundesbank, *Monthly Report of the Deutsche Bundesbank*, table VIII-7.

Series: Total cost-of-living index for all households.

Description: Monthly index.

Industrial Production

Data Source: Organization for Economic Cooperation and Development (OECD), *Main Economic Indicators*.

Series: Total industrial production.

Description: Monthly index.

Interest Rates (three-month)

Data Source: *Frankfurter Allegemeine Zeitung*.

Series: "Geldmarkt Vierteljahresgeld" in percent per annum (three-month interbank rate).

Description: Daily data.

Monetary Base (reserve-adjusted)

Data Source: Deutsche Bundesbank, *Monthly Report of the Deutsche Bundesbank*, table II-1 (components of the unadjusted monetary base) and table IV (average reserve ratio).

Series: The unadjusted base is calculated in millions of deutsche marks as total Bundesbank assets less the reserve adjustment balancing asset, foreign and domestic public authority deposits, SDR allocations, EMCF gold contributions, liquidity paper liabilities, and "other" liabilities. The reserve adjustment is made by multiplying the unadjusted base statistic by (.631 + 3.2/total average reserve ratio), where .631 = the currency percentage of the unadjusted base in the base period (January 1980), and 3.2 = (1 − .631) (the total average reserve ratio in the base period).

Description: Monthly data. Data for components of the unadjusted base refer to the last banking day of the month. The average reserve ratio is a monthly statistic.

Money Supplies

Data Source: Deutsche Bundesbank, *Monthly Report of the Deutsche Bundesbank*, table I-2.

Series: Money stock M1 and money stock M2 in millions of deutsche marks.

Description: Monthly data. Data refer to the last banking day of the month.

Adjustment: A break in the series, caused by the introduction of a new method of computation, occurs in December 1973. The 1973 statistics are adjusted using the ratio of the new to the old statistic for December 1973.

Trade Balance

Data Source: OECD, *Main Economic Indicators*.

Series: Trade balance (FOB – CIF) in billions of deutsche marks.

Description: Monthly data.

Japan

Bond Yields

Data Source: Data prior to 1981 are taken from Bank of Japan, *Economic Statistics Monthly*, table 71(2). 1981 data are taken from Planning and Research Department, Tokyo Stock Exchange, *Monthly Statistics Report*, table 8-1.

Series: Yields in percent per annum on listed government bonds (Tokyo Stock Exchange).

Description: Monthly data. Data refer to the last banking day of the month.

Consumer Prices

Data Source: Bank of Japan, *Economic Statistics Monthly*, table 119(1).

Series: General consumer price index for all Japan.

Description: Monthly index.

Industrial Production

Data Source: OECD, *Main Economic Indicators*.

Series: Total industrial production.

Description: Monthly index.

Interest Rates (three-month)

Data Source: Federal Reserve Board data base.

Series: "Over two-month ends" bill discount rate (Tokyo Stock Exchange) in percent per annum.

Description: Daily data based on Reuters quotes.

Money Supplies

Data Source: Bank of Japan, *Economic Statistics Monthly*, table 4.

Series: M1 and M2 + CD in 100 million yen.
Description: Monthly data. Data refer to the last banking day of the month.

Trade Balance
Data Source: OECD, *Main Economic Indicators.*
Series: Trade balance (FOB − CIF) in billions of yen.
Description: Monthly data.

United Kingdom

Bond Yields
Data Source: *Financial Times.*
Series: "British funds, undated, war loans 3½" in percent per annum.
Description: Daily data.

Consumer Prices
Data Source: Department of Employment, *Employment Gazette*, table 6.4.
Series: General index of retail prices, all items.
Description: Monthly index.

Industrial Production
Data Source: OECD, *Main Economic Indicators.*
Series: Total industrial production.
Description: Monthly data.

Interest Rates (three-month)
Data Source: *Financial Times.*
Series: Three-month local authority deposits (London money rates) in percent per annum.
Description: Daily data.

Monetary Base (reserve-adjusted) and Money Supplies
Data Source: Bank of England, *Quarterly Bulletin*, table 1 (monetary base components) and table II (money supplies).
Series: Money stock M1 and money stock sterling M3 in millions of pounds. The reserve-adjusted monetary base is calculated in millions of pounds as total currency in circulation plus bankers' deposits.
Description: Monthly data. Data refer to the third Wednesday of the month (second in December).

Trade balance
Data Source: OECD, *Main Economic Indicators.*
Series: Trade balance (FOB − CIF) in millions of pounds.
Description: Monthly data.

United States

With the exception of the trade balance statistics, all data are taken from the Federal Reserve Board data base. Many of these series are published in the *Federal Reserve Bulletin*, and all are available to the public.

Bond Yields
- Series: Government bonds with at least ten years to maturity.
- Description: Daily data.

Consumer Prices
- Series: Consumer Price Index.
- Description: Monthly index.

Industrial Production
- Series: Total industrial production.
- Description: Monthly index.

Interest Rates (three-month)
- Series: Treasury bill rates.
- Description: Daily data.

Monetary Base (reserve-adjusted) and Money Supplies
- Series: Reserve-adjusted monetary base, $M1-B$, $M2$, and $M3$.
- Description: Weekly Wednesday data.

Trade Balance (data through 1978)
- Data Source: Department of Commerce, *Highlights of U.S. Export and Import Trade*, Exports table E-1; Imports table I-1.
- Series: Domestic and foreign exports, excluding Department of Defense shipments, in millions of dollars on a FAS value basis; general imports in millions of dollars on a customs valuation basis changing to a FAS basis in 1974.
- Description: Monthly data.
- Adjustment: 1973 statistics are adjusted to a FAS value basis using the 1974 average ratio of customs valuation to FAS value.

Trade Balance (1979–81 data)
- Data Source: Department of Commerce, *Summary of U.S. Export and Import Merchandise Trade*, December 1980 (advance statistics for *Highlights of U.S. Export and Import Trade*), Exports table 3; Imports table 5.
- Series: Total domestic exports, excluding Department of Defense grant-aid, in millions of dollars on a FAS value basis; general imports in millions of dollars on a FAS value basis.
- Description: Monthly data.

References

Barro, Robert J. 1976. Integral constraints and aggregation in an inventory model of money demand. *Journal of Finance* 31:77–88.

Baumol, William. 1952. The transactions demand for cash—An inventory-theoretic approach. *Quarterly Journal of Economics* 66:545–56.

Bilson, John F. O. 1978. Rational expectations and the exchange rate. In *The economics of exchange rates*, ed. J. Frenkel and H. Johnson. Reading, Pa.: Addison-Wesley.

———. 1979. The deutsche mark/dollar rate—A monetary analysis. In *Policies for employment, prices, and exchange rates*, ed. K. Brunner and A. Meltzer. Carnegie-Rochester Conference Series on Public Policy, vol. 11. Amsterdam: North-Holland.

———. 1981. The "speculative efficiency" hypothesis. *Journal of Business* 54:435–51.

Chow, Gregory C. 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28:591–605.

Cumby, Robert, and Maurice Obstfeld. 1981. A note on exchange rate expectations and nominal interest differentials. *Journal of Finance* 36:697–703.

Dornbusch, Rudiger. 1976*a*. The theory of flexible exchange rate regimes and macroeconomic policy. *Scandinavian Journal of Economics* 78:255–75.

———. 1976*b*. Expectations and exchange rate dynamics. *Journal of Political Economy* 84:1161–76.

Driskill, Robert A., and Steven M. Sheffrin. 1981. On the mark: Comment. *American Economic Review* 71:1068–74.

Fischer, Stanley. 1982. Relative price variability and inflation in the United States and Germany. *European Economic Review* 18:171–96.

Frankel, Jeffrey A. 1979. On the mark: A theory of floating exchange rates based on real interest differentials. *American Economic Review* 69:610–22.

———. 1981. On the mark: Reply. *American Economic Review* 71: 1075–82.

Frenkel, Jacob A. 1976. A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence. *Scandinavian Journal of Economics* 78:200–24.

———. 1981*a*. The collapse of purchasing power parities during the 1970's. *European Economic Review* 16:145–65.

———. 1981*b*. Flexible exchange rates, prices, and the role of news: Lessons from the 1970's. *Journal of Political Economy* 89:665–705.

Genberg, Hans. 1978. Purchasing power parity under fixed and flexible exchange rates. *Journal of International Economics* 8:247–76.

Geweke, John, and Edward Feige. 1979. Some joint tests of the efficiency of markets for forward foreign exchange. *Review of Economics and Statistics* 61:334–41.

Glaessner, Thomas. 1982. Theoretical and empirical essays on spot and forward exchange rate determination. Ph.D. diss., University of Virginia.

Goldfeld, Stephen M. 1973. The demand for money revisited. *Brookings Papers on Economic Activity* 3:577–638.

———. 1976. The case of the missing money. *Brookings Papers on Economic Activity* 3:683–739.

Goldfeld, Stephen M., and Richard Quandt. 1965. Some tests for homoscedasticity. *Journal of the American Statistical Association* 60:539–47.

Hacche, Graham, and John Townend. 1981. Exchange rates and monetary policy: Modeling sterling's effective exchange rate, 1972–80. In *The money supply and the exchange rate*, ed. W. Eltis and P. Sinclair. Oxford: Oxford University Press.

Hakkio, Craig S. 1981. Expectations and the forward exchange rate. *International Economic Review* 22:663–78.

Hansen, Lars P., and Robert J. Hodrick. 1980*a*. Forward exchange rates as optimal predictions of future spot rates: An econometric analysis. *Journal of Political Economy* 88:829–53.

———. 1980*b*. Speculation and forward foreign exchange rates: Some tests of cross-currency restrictions. Carnegie-Mellon University. Mimeo.

Haynes, Stephen E., and Joe A. Stone. 1981. On the mark: Comment. *American Economic Review* 71:1060–67.

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* 67:942–48.

Litterman, Robert B. 1979. Techniques of forecasting using vector autoregressions. Federal Reserve Bank of Minneapolis, Working Paper no. 115.

Meese, Richard A., and Kenneth S. Rogoff. 1983. Empirical exchange rate models of the seventies: Are any fit to survive? *Journal of International Economics* 14:3–24.

Meese, Richard A., and Kenneth J. Singleton. 1982. On unit roots and the empirical modeling of exchange rates. *Journal of Finance* 37: 1029–35.

Miller, Merton, and Daniel Orr. 1966. A model of the demand for money by firms. *Quarterly Journal of Economics* 80:413–35.

Parzen, Emanuel. 1975. Multiple time series: Determining the order of approximating autoregressive schemes. Technical Report no. 23. State University of New York at Buffalo.

Simpson, Thomas D., and Richard D. Porter. 1980. Some issues involving the definition and interpretation of the monetary aggregates. In *Controlling the monetary aggregates III*. Federal Reserve Bank of Boston Conference Series no. 23.

Sims, Christopher A. 1980. Macroeconomics and reality. *Econometrica* 48:1–48.

Tobin, James. 1956. The interest elasticity of transactions demand for cash. *Review of Economics and Statistics* 38:241–47.

Tryon, Ralph. 1979. Testing for rational expectations in foreign exchange markets. Federal Reserve Board, International Finance Discussion Paper no. 139.
Westerfield, Janice M. 1977. An examination of foreign exchange risk under fixed and floating rate regimes. *Journal of International Economics* 7:181–200.
Whalen, Edward L. 1966. A rationalization of the precautionary demand for cash. *Quarterly Journal of Economics* 80:314–24.

Comment Nasser Saïdi

The Meese and Rogoff paper is an assessment and adds up to an indictment of the large body of empirical models elaborated to interpret the time series behavior of exchange rates in the 1970s. This is an introspective, examining-one's-own-belly-button, piece of work. Not having placed any bets on the horses running in the Meese-Rogoff races, I admit to being in basic agreement with their dim assessment of existing empirical models.

I discuss, first, the contributions and substantiated conclusions of the paper. This is followed by some comments on the methodology, including discussion of some technical issues. Finally, I expand on the reasons advanced by Meese and Rogoff (hereafter, M-R) for the relative failure of the structural models to accurately forecast exchange rates.

Contributions and Conclusions

The main contributions of this paper are:

a. It evaluates the forecast accuracy of a number of competing models on common ground. Estimates and forecasts are generated for identical sample periods, using common, point-in-time data and common estimation methods. The models are evaluated on common ground by avoiding noncomparability problems caused by sampling differences between models.

b. A more stringent criterion is used for evaluating the models' post-sample forecast accuracy. This is important since there is no guarantee that a model with "good" in-sample operating charcteristics will perform well out of sample.

c. The estimation period covers the major portion of the 1970s experience with floating rates, and the sampled exchange rates display substantial variance. The data should be informative and should discriminate between alternative models.

Nasser Saïdi is a professor of economics at the Graduate Institute of International Studies in Geneva, Switzerland.

This paper addresses both the Meese and Rogoff paper in this volume and Meese and Rogoff (1981).

What conclusions can be drawn from M-R's work, and what are the implications?

1. For all exchange rates, for nearly all forecast horizons, and for alternative measures (RMSE, MAE) of out-of-sample forecast accuracy, a simple random walk (SRW) model is a more accurate forecaster of the level of the (log) exchange rate than any structural model or unconstrained VAR.

2. Although a SRW model ranks best, it is not a good predictor of exchange rates. The correlations between actual and predicted outcomes are low. The SRW forecast is of poor quality.

3. Structural models fail, rather miserably, at both predicting and explaining the postsample behavior of exchange rates. They fail to *explain* in that, provided with the actual realizations of the "exogenous" variables (which in principle introduce a bias in favor of the structural models), their predictions of realized exchange rates are worse than that of a SRW model.

4. A tentative conclusion is that the structural models fail because their parameter estimates are unstable over time. If coefficient estimates that provide good in-sample fits are combined with the postsample values of the exogenous variables, the structural models do not explain exchange rates.

5. However, even if one partially resolves the issue of coefficient uncertainty by generating forecasts using a range of values consistent with theory—which amounts to giving a weight of unity to prior information—and the postsample realizations of exogenous variables, the structural models still fail to predict exchange rates.

6. Consider the relative performance of nonstrict time series models (i.e., exclude the SRW and univariate AR which only use the past history of the exchange rate). It turns out that an unrestricted VAR does better—in terms of average rank for each specific measure and forecast horizon—than any of the structural models. Since the various structural models are restricted versions of the VAR, the sample evidence rejects the joint exogeneity and distributed lag restrictions implicit in the structural models.

Most of these results should come as no surprise. The work of Nelson (1972) and Fair (1979) reaches similar conclusions with respect to the forecast accuracy of large-scale macroeconometric models as compared to time series models. M-R demonstrate that the poor performance of macroeconometric exchange rate models cannot be accounted for by uncertainty due to exogenous variable forecasts or coefficient estimates.

Methodology and Evaluation

I have a number of technical quibbles with the M-R methodology and evaluation criteria. First, it is highly likely that the residual errors for different exchange rates are contemporaneously correlated. Hence, a joint estimation strategy that utilizes the information available in the covariance matrix would provide more efficient coefficient estimates and potentially improve the forecast accuracy of the structural models. In addition, to the extent that the forecast errors for different exchange rates are correlated, the relative failure of the structural models for each exchange rate does not provide additional, independent evidence. Second, when estimated in the levels, the structural models typically yield serially correlated residuals. Fair's (1979) results indicate that appropriate modeling of the time series error structure can vastly improve the forecast accuracy and performance of structural models. M-R's handling of this problem is rather cavalier. The modeling of the residual is restricted on an a priori basis to a first-order AR, with no use made of sample information. Appropriate modeling of the residuals using the sample evidence may improve the relative performance of the structural models. Third, as M-R note, forecasts for more than one period ahead follow a moving average process. Hence, the only statistically meaningful tests for evaluating forecasts obtained from different models are those for the one period ahead forecasts. The summary measures for multiperiod forecast horizons are not very informative. For the purposes of comparing models, it would be as informative to consider the population correlation coefficients between actual and predicted exchange rates and provide a ranking.

Despite these reservations, I doubt that the major M-R conclusions would be reversed by using more sophisticated estimation procedures. We are left with the somewhat disquieting result that existing structural models that seek to interpret exchange rate time series by appealing to the behavior of observable macroeconomic variables fail to explain and forecast. However, the observation that a SRW model provides better forecasts does not provide an interpretation of exchange rate movements in the 1970s.

Reasons for Failure

It seems clear that the structural models fail to explain and forecast exchange rates because they are misspecified. M-R suggest—but provide no evidence in their paper to substantiate—that the major problem lies in the instability of the money demand specification. The suggestion does not seem unreasonable, given the recent accumulating evidence on the temporal instability of estimated money demand functions or, more generally, of behavioral relations in the monetary sector. However, the suggestion is moot unless the potential factors leading to misspecification and for instability of estimated reduced forms are identified. One, in my view, crucial source of misspecification is the inadequate modeling of expectations formation in existing structural models. In particular, the role of the distinction between anticipated and unanticipated movements in the exogenous driving variables is vastly underplayed. Asset pricing models of exchange rate determination incorporating rational expectations typically lead to a specification such as:

$$S_t = \sum_{j=0}^{\infty} A_j E_{t+j-1} Z_{t+j} + B(Z_t - E_{t-1} Z_t) + u_t, \tag{1}$$

where S is the spot exchange rate, Z stands for a vector of exogenous variables, E is the expectations operator, and u_t is interpreted as being generated by differences in the information available to economic agents pricing foreign exchange and the econometrician (see Sargent 1981). The forecast error is given by:

$$S_t - E_{t-1} S_t = B(Z_t - E_{t-1} Z_t) + (u_t - E_{t-1} u_t),$$

proportional to the unanticipated movements in Z_t and in u_t (although some specifications of u_t might lead to the restriction $E_{t-1} u_t = 0$). As an example, equation (1) would say that fully anticipated, permanent movements in the money stock (a component of Z) would lead to one-to-one changes in the exchange rate (neutrality), but unexpected changes would have an effect that depends on the magnitude and temporal characteristics of the process generating the money stock.

It is clear from equation (1) that if the distinction between anticipated and unanticipated movements in Z is important, then models of the form $S_t = CZ_t + V_t$ are unlikely to provide good out-of-sample forecasts and will not adequately explain exchange rate movements. If (1) is an appropriate model, then it is not surprising that the structural models fail to explain and forecast even when provided with the actual, realized values of the exogenous variables.

Equation (1) suggests some further reasons for the potentially poor performance of the structural models and the unconstrained VAR. The first reason is the Lucas critique argument: Perceived changes in the processes or policies generating the Z's will lead to changes in the A's and B's of equation (1). The second is a timing or dating of variables problem, which is intimately linked to the information set available to economic agents. If one uses, as M-R do, point-in-time exchange rate data, then (1) suggests that the Z variables should be appropriately dated. For example, end-of-month exchange rates may well appear unrelated to end-of-month money supplies if information on the latter is known only with a two-week lag. This is a case in which information subsequently available to the econometrician was not contemporaneously observable to agents. The question of the availability and timing of information, including announcement-type effects (see Schwert 1981), is neglected in the M-R paper as well as in the empirical literature they assess. Finally, and related to the Lucas critique, there is the "peso problem" and the "finance minister problem." The peso problem is one in which an event (say a change in monetary policy such as the formation of the European Monetary System (EMS)) is expected to occur at some future date. This alters the exchange rate path in anticipation of the event. Both a structural model that inadequately models expectations and a SRW model will fail to explain and forecast exchange rates. This problem is amenable to correction to the extent that anticipations can be conditioned on the past and contemporaneous movement in observable exogenous variables (for a good example of how to proceed, see LaHaye's 1981 analysis of anticipations of currency reforms). On the other hand, the finance minister problem is one in which an event is expected to occur, exchange rates respond, but the event does not happen (the particular finance minister is not appointed). In this case, exchange rate expectations will appear unrelated to the past, contemporaneous, or subsequent evolution of the included exogenous variables. While the finance minister problem leads to a theory of the error term u_t in equation (1), it is not clear how the estimation or modeling should proceed.

To conclude, Meese and Rogoff have provided a useful and sorely needed assessment of existing empirical exchange rate models. The estimated VARs suggest that the exogeneity restrictions and lag distributions of the structural models are not credible and are rejected by the sample evidence. One direction for research would be to take account of the information contained in the VARs in modeling exchange rates. More important, the inappropriate modeling of expectations appears to be the main reason for the failure of the structural models to explain and forecast exchange rates in the 1970s.

References

Fair, Ray. 1979. An analysis of the accuracy of four macroeconometric models. *Journal of Political Economy* 87:701–18.

LaHaye, Laura. 1981. Inflation and currency reform. University of Illinois at Chicago Circle. Mimeo.

Meese, Richard, and Kenneth Rogoff. 1981. Empirical exchange rate models of the seventies: Are any fit to survive? Federal Reserve Board, International Finance Discussion Paper no. 184.

Nelson, Charles. 1972. The prediction performance of the FRN-MIT-PENN model of the U.S. economy. *American Economic Review* 62:902–17.

Sargent, Thomas. 1981. Interpreting economic time series. *Journal of Political Economy* 89:213–48.

Schwert, William. 1981. The adjustment of stock prices to information about inflation. *Journal of Finance* 36:15–29.

Comment Michael K. Salemi

The present contribution of Meese and Rogoff (hereafter, MR) and Meese and Rogoff (1981) constitute a comprehensive appraisal of the empirical validity of several asset approach models of the spot price of foreign exchange. MR test spot rate equations implied by three models: the flexible-price monetary model of Frenkel and Bilson, the sticky-price monetary model of Dornbusch and Frankel, and Hooper and Morton's sticky-price asset model which incorporates the current account. The authors find that these models are not correctly specified structural equations. They present abundant and compelling evidence to support this view.

MR test the so-called structural models in three different ways. In Meese and Rogoff (1981), they perform a battery of tests that compare the out-of-sample predictions of the models with those of a naive alternative, a random walk model for the spot rate. They estimate restricted versions of their equation (2) using monthly data beginning at March 1973 and generate out-of-sample predictions at one-, three-, six-, and twelve-month horizons for two periods: December 1976 through November 1980, and December 1978 through November 1980. The predictions of each model use actual realized values of the appropriate explanatory variables. MR find that the predictions of the random walk model are consistently more accurate than those of any of the structural models. Their result apears robust across foreign currencies, the test period chosen, the statistic used to measure the size of prediction errors, and the prediction horizon.

Although MR refer to their work as a forecasting experiment, their results are solid evidence to reject the models themselves. The parameters of a correctly specified structural equation are invariant with respect to regimes. By their procedure, MR confront equation (2) and the models it characterizes with a regime shift—a set of realizations for the explanatory variables different than that used to estimate the parameters. The failure of these models to explain the spot rate as well as the random walk model is evidence that the parameters of equation (2) are not invariant to the regime shift.

In the present paper, MR subject the asset models to additional tests. They first estimate a system of vector autoregressions (VAR) for the seven variables that appear in the Hooper-Morton model, the least restrictive of the models they study. Two findings are particularly noteworthy. First, none of the variables in the system is exogenous in the sense of Granger and Sims. The econometric implication of this result is

Michael K. Salemi is associate professor in the Department of Economics at the University of North Carolina at Chapel Hill.

that, for the period under study, the parameters of the asset models are not consistently estimated by ordinary least squares. Of course, this finding also challenges model builders to explain the mutual endogeneity of these variables.

MR reach their second interesting finding by inverting the estimated VAR to obtain a moving average representation for the variables in the system. After a necessary preordering of contemporaneous influences, MR compute a variance decomposition for each variable in the system. At short horizons (especially one and three months) they find that most of the variation in the spot rate prediction error is attributable to the innovation in the spot rate itself rather than to the innovations of the other variables in the system. At longer horizons, however, innovations in the other variables account for as much as 54 percent of the prediction error variance in the spot rate. Unfortunately, no consistent ranking of the importance of the other variables emerges. Relative output levels and short-term interest rates appear relatively unimportant. Innovations in the trade balance seem important for the yen, while relative money stocks and the long-term interest rate differential appear more important for the mark.

Finally, MR demonstrate that the results of their first paper (1981) are not likely to be merely the result of imprecise estimation of the asset model coefficients. On the basis of the money demand and monetary trade literatures, they define a grid of probable values for the parameters of the asset models. For each point on the grid they generate spot rate predictions for the entire period between March 1973 and June 1981. For horizons less than or equal to twelve months, virtually no parameterization of the asset approach models predicts as well as the random walk model. But for longer horizons, the best parameter configuration of each of the three models predicts the spot rate substantially more accurately than the random walk model.

There are two ways to view this result. Because the grid search is an in-sample procedure, one might by reasonably reluctant to conclude that the spot rate models studied have long-run validity. On the other hand, I find it an interesting possibility that, in the short run, the spot rate behaves like the price of a speculative asset but that over longer horizons its equilibrium value is systematically related to other economic variables, as the asset models predict.

In conclusion, Meese and Rogoff have, with exceptional skill and concern for detail, cast serious doubt on the validity of the Frenkel-Bilson, Dornbusch-Frankel, and Hooper-Morton models. The sharp contrast between the in-sample and out-of-sample performance of these models cautions us to exercise more care in testing the models we estimate. Finally, it bears pointing out that the results of Meese and Rogoff are not evidence to reject the asset approach itself. That approach, in its most general form, says simply that the spot rate adjusts to guarantee that portfolio owners are satisfied with the foreign asset component of their portfolios.

References

Meese, Richard, and Kenneth Rogoff. 1981. Empirical exchange rate models of the seventies: Are any fit to survive? Federal Reserve Board, International Finance Discussion Paper no. 184.