ELSEVIER

Journal of International Money and Finance
17 (1998) 813–830

Journal of International Money and Finance

Integration, cointegration and the forecast consistency of structural exchange rate models

Y.-W. Cheung, M.D. Chinn*

Department of Economics, University of California, Santa Cruz, CA 95064, USA

Abstract

We propose an alternative set of criteria for evaluating forecast rationality: the forecast and the actual series (1) have the same order of integration, (2) are cointegrated and (3) have a cointegrating vector consistent with long-run unitary elasticity of expectations. We denote forecasts that meet these criteria as 'consistent'. Forecasts generated from monetary models generally pass (1). However, using the Johansen procedure, cointegration fails to hold the longer the horizon. Of the cointegrated pairs, (3) is not generally rejected. Using the Horvath–Watson procedure, imposing the unitary coefficient restriction, we find fewer instances of consistency, although a higher proportion of the cases of consistency are found at the longer horizons. © 1998 Elsevier Science Ltd. All rights reserved.

JEL classifications: F31, C50

Keywords: Forecasts; Exchange rate; Cointegration

1. Introduction

Numerous studies have compared the forecasting performance of various exchange rate models, structural and non-structural, against that of the random walk model. Results from these studies tend to corroborate the finding that it is extremely difficult to out-predict a random walk using structural or other time

* Corresponding author. Tel.: +1 408 4592079; fax: +1 408 4595900; e-mail: chinn@cats.ucsc.edu.

PII: S0261-5606(98)00009-6

series models. This result has held up for a wide variety of forecast metrics, structural and time series models, estimation techniques and sample periods.[1]

This study attempts to evaluate forecasts from structural models based on the time series properties of these forecasts. Instead of examining the commonly used measures of forecast accuracy, such as the mean squared error, mean absolute deviation and the serial correlation of the forecast errors, we explore some basic time series properties of forecasts. In particular, we examine whether forecasts from structural models and the spot exchange rate series (1) have the same order of integration, (2) are cointegrated and (3) have a cointegrating vector consistent with long run unitary elasticity of expectations.

The first property relates to the persistence of forecasts and spot exchange rates, as measured by the order of integration. The other two properties are related to how exchange rates and their respective forecasts are related in the long-run. While exchange rate forecasts may deviate from the observed exchange rates in the short-run, we expect a forecast of any practical relevance should have the above properties. We label the condition where these three properties hold as the 'consistency' of a forecast.[2] That is, a forecast is consistent if it has a one-to-one relationship with the spot exchange rate in the long-run.[3] This concept involves the behavior of the forecast relative to the actual, *over time* and *on average*.

At this juncture, it may be worthwhile to discuss in detail several reasons why this notion of forecast consistency is useful and necessary, given the plethora of competing criteria. First, the notion of consistency focuses on the long-run property of forecasts, and hence is weaker than the one conventionally used in evaluating forecast rationality. It does not, for example, impose any further restrictions on the forecast errors, above and beyond the requirement that they be weakly covariance stationary.[4] In this approach we *test* whether this condition holds, rather than merely assuming it does, which is typically done when comparing, for instance, root mean squared errors (RMSEs).

Second, in our approach, a forecast can meet the requirement of consistency and, at the same time, its errors need not be serially uncorrelated. This can happen, for example, when the correct model is used but the data on the fundamentals are contaminated by stationary measurement errors. Such a situation is *very* likely to occur in the case of typical asset-based models which incorporate

[1]Some recent attempts to overturn the Meese and Rogoff (1983) results include Cheung (1993) (fractional integration models), Diebold and Nason (1990) (autoregressive non-parametric methods), Meese and Rose (1991), (non-linear models) and Chinn and Meese (1995) (non-linear functions, structural models and long horizons).

[2]The usage of 'consistency' here is different from that in econometrics, where it denotes convergence in probability, a concept that involves the property of the estimator when the sample size approaches infinity. It also differs from a recent definition attributable to Froot and Ito (1989).

[3]Fischer (1989) and Liu and Maddala (1992) apply the concepts of integration and cointegration to testing for relationships between the survey-based forecasts and the actual series. Fischer (1989) does so in the context of the US money stock, while Liu and Maddala (1992) address exchange rates.

[4]In the literature, a forecast is said to be 'rational' if the forecast errors have a zero mean *and* zero serial correlation.

information on industrial production, money stocks and price indices. Thus, the consistency requirement represents a more realistic way to evaluate exchange rate forecasts from structural models.

Third, rejection of the unitary elasticity criterion may arise for reasons unrelated to irrationality. Consider the situation where the right-hand side variables used to forecast the exchange rate are measured with error. In this case, the unitary elasticity restriction might be violated even though the forecasts are in some sense optimal. An example of how measurement error can induce deviations from unitary elasticities in a purchasing power parity cointegrating vector is provided in Cheung and Lai (1993b).

Although this definition of consistency is weaker than those typically employed in the forecast evaluation literature, we find that it is relatively difficult for the forecasts generated by the structural models to fulfill all three criteria in this sample. This outcome suggests that the condition proposed is useful in discriminating between different forecasts.

2. Exchange rate models: estimation and forecasting

2.1. Exchange rate models

This study examines the consistency property of forecasts from three monetary models: the Frenkel (1976) flexible price model; the Frankel (1979) sticky price model; and the Dornbusch (1976) tradables-non-tradables model. All these models start with conventional money demand functions for both the domestic and foreign economies, and impose the condition that expected depreciation equals the nominal interest differential plus an exogenous risk premium on domestic assets that may or may not be zero. These models can be written, respectively, as:

$$\text{Model 1: } s = (m - m^*) - \phi(y - y^*) + \mu(i - i^*), \tag{1}$$

$$\begin{aligned}\text{Model 2: } s = {} & (m - m^*) - \phi(y - y^*) + (\mu + 1/\theta)(\pi - \pi^*) \\ & -(1/\theta)(i - i^*),\end{aligned} \tag{2}$$

$$\begin{aligned}\text{Model 3: } s = {} & (m - m^*) - \phi(y - y^*) + (\mu + 1/\theta)(\pi - \pi^*) \\ & -(1/\theta)(i - i^*) + \beta q \\ \equiv {} & [(p^T - p^N) - (p^{T*} - p^{N*})],\end{aligned} \tag{3}$$

where s, m, y and q are the logarithms of the exchange rate (domestic currency per unit of foreign currency), money supply, real income and the intercountry relative price of tradables to non-tradables (p^T–p^N), and i and π are the levels of the nominal interest and inflation rates, respectively. An asterisk denotes a foreign variable.

Model 1 contains only the terms in monies, incomes and nominal interest rates, and relies on the assumption that purchasing power parity (PPP) holds continuously. This 'flexible price' monetary model subsumes the Lucas (1982) model since the latter model contains monies and real incomes but no interest rate term. Model 2, a 'sticky price' monetary model slow adjustment of goods prices (at rate θ) and instantaneous adjustment of asset prices, thus yielding the well-known overshooting characteristic. Model 3 is motivated by the failure of PPP to hold for broad price indices, such as the consumer price index and GDP deflator. One approach is to make an explicit recognition of non-traded goods, and to posit that PPP only holds for tradable goods. If the aggregate price level index is represented by a Cobb–Douglas function of the individual non-traded and traded price indices (with weight β on non-tradables) then model 3 is obtained.

2.2. Estimation

Since it is generally accepted that exchange rates and their fundamentals are well approximated by unit root processes, we will estimate all three of these models in first difference form, using OLS and two-stage least squares (2SLS) procedures. An instrumental variable approach such as 2SLS is appropriate because the right hand side variables — such as interest rates and money stocks — can plausibly be interpreted as being jointly determined with the exchange rate.[5]

In addition to the first-difference specification, we also implement the error correction version of these models. The error correction model (ECM) variants include the error correction term (to be discussed below) lagged once, and the first difference of fundamentals lagged once. Thus all regressors in the ECM models are predetermined, and 1 month ahead forecasts are true ex-ante forecasts.

The Chinn and Meese (1995) methodology is used to construct the error correction term that captures the long-run relationship between exchange rates and their fundamentals. We assume that log linear versions of Eqs. (1)–(3) are appropriate in the long run, and impose a set of coefficient restrictions for each of the models. For all models, the money supply and income elasticities are the same (unity and 0.75, respectively). The coefficients on interest rates, inflation rates and relative prices vary by model, although the coefficients on the first two variables are functions of the interest rate semi-elasticity, which we assume is 4.5. The goods market speed of adjustment parameter is taken to be 0.5 on an annual basis; this corresponds to deviations from PPP damping at rates 0.94 for monthly data. The final parameter of interest is the share of non-tradables in the aggregate price index, β, which we take to be 0.5.

These estimated models have explanatory power similar to, or greater than, those reported in the exchange rate forecasting literature. This method of imposing prior values on the error correction term reflects the consensus view that the post-Bretton Woods period is too short to obtain reliable estimates of the long-run coefficients. In fact, in-sample estimates have rarely correlated with out-of-sample success (see Frankel and Rose, 1995).

[5]Assuming rational expectations, appropriate instrumental variables include elements in the information set such as lagged variables. We use lags 2–4 of the right hand side variables, since there is evidence of MA1 serial correlation in the first difference specifications.

2.3. The forecasting exercise

We evaluate the out-of-sample explanatory power of our representative models over two forecast periods. Our choice of forecast periods is arbitrary; the first starts with the end of the recession in the US in 1982, and the second corresponds to the period after the Louvre Accord in April 1987.

In the experiments reported below, the original estimation period for the first sample is June 1973 through December 1982 (115 observations). We then 'roll' through our sample ending in August 1993 to produce 128, 123 and 117 1-, 6- and 12-months ahead forecasts, respectively. Whenever necessary, forecasts use actual realized values of the right hand side variables. As we 'roll' through each forecast period, parameter estimates are updated with the addition of each new data point. The original estimation period for the second sample is June 1973 to June 1987 (169 observations). We then perform an analogous 'rolling regression' procedure, to produce 74, 69 and 63 1-, 6- and 12-month ahead forecasts.

3. Unit root test and cointegration analysis

For a time series $\{y_t\}$, $t = 1,...,t'$, the ADF unit root test is based on the regression

$$\Delta y_t = c + \mu t + \gamma_0 y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t. \tag{4}$$

Δ is the differencing operator defined by $\Delta y_t \equiv y_t - y_{t-1}$. The following procedure is used to determine the lag order parameter k. First, the Akaike Information Criterion (AIC) are used to select the lag order among specifications $k = 1,..., 13$. Then, residuals from the selected specification are tested for serial correlation. If significant serial correlation is detected, the lag length is increased until the model passes the residual test.

Testing for cointegration proceeds in the following manner. Consider in general an $m \times 1$ vector $\mathbf{x}_t$ of I(1) variables and its VAR(p) representation:

$$\Delta x_t = \mu + \Gamma_1 \Delta x_{t-1} + \Gamma_2 \Delta x_{t-2} + ...\Gamma_{p-1} \Delta x_{t-p+1} + \Pi x_{t-p} + u_t, \quad \Pi = \alpha\beta' \tag{5}$$

where Γ_1, Γ_2,... Γ_{p-1}, Π are $m \times m$ matrices of unknown parameters. α and β are $m \times r$ matrices, representing the rate of reversion and cointegrating parameters, respectively. See Johansen (1991) for a more detailed account of this cointegration methodology.

Johansen proposes two tests for inferring the number of cointegrating vectors. The trace statistic is used for testing the null hypothesis of at most r cointegrating vectors against the alternative of m cointegrating vectors. The maximal eigenvalue statistic is used in testing the null hypothesis of $r-1$ against r cointegrating vectors. According to our definition of 'consistency', forecasts should be cointegrated with the actual series. Failing this, forecasts could drift infinitely far away from the actual series.

A stronger requirement for the consistency of a forecast is that the coefficients in the cointegrating vector are (1−1). Following Johansen and Juselius (1990) and Johansen (1991), the hypothesis of a linear constraint on the cointegrating vector can be expressed as:

$$H_G\colon \beta = GB, \tag{6}$$

where G is a known $m \times r_0$ matrix of full rank r_0 and B is a $r_0 \times r$ matrix of unknown parameters ($m \geq r_0 \geq r$). If $r_0 = r$, the cointegrating space is fully specified. If $r_0 = m$, then no restriction is imposed on β. Note that G is the matrix that defines the coefficient restriction. In terms of (6), the unitary elasticity restriction is described by $(1-1)'$, so $r_0 = 1$ in this case. In the following section, the Johansen (1991) likelihood ratio test statistic will be used to evaluate H_G.

4. Estimation results

The Japanese yen, German mark and Canadian dollar exchange rates against the US dollar are considered. Monthly data on exchange rates and the related fundamentals were obtained from the OECD *Main Economic Indicators* database. Details are provided in the Data Appendix.

4.1. Unit root test results

In accord with previous research, we find that we cannot reject the null of unit roots in all the actual nominal exchange rate series using the 5% marginal significance level. The unit root test results for the forecasts are presented in Table 1. Detailed test results are not reported given space constraints; they can be found in Cheung and Chinn (1997). We find that all the forecast series in the longer post-1982 sample also appear integrated. These results therefore fulfill the first condition of consistent forecasts — that the series share the same order of integration. However, for the shorter post-Louvre sample, several yen/dollar forecasts reject the unit root null. Since the outcome is a rejection of the null hypothesis, this result cannot be attributed to the low power of the unit root tests. Nor can the source of this result be located in the specific estimation technique — both OLS and two stage least squares specifications appear to be trend stationary at 1-month (6-month for OLS) or all horizons (2SLS), across a variety of models. Hence, it appears that the peculiarity is specific to the forecasts of the yen/dollar for this shorter forecasting period.

Table 1
Augmented Dickey–Fuller test results

	Model	Sample 1 January 1983–August 1993			Sample 2 June 1987–August 1993		
		1-month ahead	6-month ahead	12-month ahead	1-month ahead	6-month ahead	12-month ahead
Differences							
Canadian $/US$	1	−1.286	−1.390	−1.706	−1.282	−1.219	−1.380
	2	−1.029	−1.409	−1.609	−1.305	−1.279	−1.451
	3	−1.036	−1.394	−1.580	−1.297	−1.292	−1.473
Mark/US$	1	−1.731	−1.870	−2.148	−2.732	−2.466	−2.067
	2	−1.754	−1.888	−2.214	−2.277	−2.156	−2.391
	3	−1.929	−1.963	−2.433	−2.261	−2.140	−2.449
Yen/US$	1	−1.707	−1.828	−1.615	−2.510	−3.486*	−3.414
	2	−1.696	−1.842	−1.619	−2.482	−3.512*	−3.463
	3	−1.704	−1.836	−1.594	−2.475	−2.661	−3.359
Error correction							
Canadian $/US$	1	−1.108	−1.426	−1.525	−1.184	−1.402	−1.824
	2	−1.147	−1.473	−1.488	−1.192	−1.475	−1.755
	3	−1.154	−1.516	−1.450	−1.218	−1.530	−1.726
Mark/US$	1	−1.789	−2.211	−2.283	−2.082	−2.831	−1.519
	2	−1.691	−2.200	−2.321	−2.813	−2.711	−1.586
	3	−1.695	−2.156	−2.361	−2.809	−2.693	−1.764
Yen/US$	1	−1.628	−1.762	−1.368	−1.802	−2.933	−2.747
	2	−1.589	−1.801	−1.492	−1.825	−3.043	−2.414
	3	−1.598	−1.819	−1.463	−1.867	−3.063	−2.556
2SLS							
Canadian$/US$	1	−1.206	−1.249	−1.627	−1.198	−1.238	−1.643
	2	−1.195	−1.665	−1.866	−1.120	−1.386	−1.748
	3	−1.225	−1.978	−2.073	−1.249	−1.943	−2.598
Mark/US$	1	−2.010	−2.164	−3.261	−2.856	−2.669	−2.697
	2	−1.648	−2.197	−2.743	−2.843	−3.164	−2.556
	3	−1.547	−1.792	−3.031	−2.914	−2.971	−2.285
Yen/US$	1	−2.248	−2.005	−2.313	−3.905*	−5.771*	−5.428*
	2	−2.453	−2.240	−2.145	−3.485*	−3.416	−4.930*
	3	−2.332	−2.303	−2.341	−3.492*	−3.517*	−5.027*

Notes: ADF statistics for regressions selected by AIC. *, indicates significance at 5% MSL using Cheung and Lai (1995) finite sample critical values. Model 1 is the flexible-price monetary model (Frenkel, 1976), Model 2 is the sticky-price monetary model (Frankel, 1979), and Model 3 is the monetary model incorporating relative non-tradables prices. See Appendix 1, Cheung and Chinn (1997) for details of the unit root tests.

4.2. Cointegration test results

The results of applying the Johansen cointegration maximal eigenvalue test to spot exchange rates and forecasts are reported in Table 2. We applied the

Table 2
Cointegration Test Results (Sample 1: January 1983—August 1993)

	Model	Forecasting horizons					
		1-month ahead		6-month ahead		12-month ahead	
		$r = 0$	$r = 1$	$r = 0$	$r = 1$	$r = 0$	$r = 1$
Differences							
Canadian$/US$	1	88.938*	0.905	16.692*	9.178	32.142*	3.919
	2	76.369*	0.856	66.034*	4.582	27.915*	4.134
	3	105.561*	0.991	64.554*	4.324	26.554*	3.734
Mark/US$	1	64.490*	1.865	55.108*	1.702	8.310	3.972
	2	38.244*	1.407	55.135*	1.736	8.216	4.128
	3	36.957*	1.393	15.128	2.272	8.672	3.956
Yen/US$	1	82.985*	0.433	52.992*	0.684	11.901	1.297
	2	81.983*	0.439	51.685*	0.681	11.635	1.314
	3	80.412*	0.432	51.133*	0.693	12.009	1.273
Error correction							
Canadian$/US$	1	19.764*	1.096	14.050	4.563	28.289*	3.106
	2	17.090	1.132	66.998*	3.290	25.560*	2.938
	3	66.001*	0.893	69.886*	3.317	24.584*	3.159
Mark/US$	1	52.433*	1.999	55.370*	1.804	5.885	4.297
	2	45.954*	1.720	55.713*	1.768	6.960	4.786
	3	44.564*	1.822	54.638*	1.815	6.774	4.840
Yen/US$	1	104.395*	0.450	59.806*	0.781	13.010	1.071
	2	98.589*	0.464	67.235*	0.798	13.743	1.244
	3	100.130*	0.452	62.678*	0.800	12.901	1.218
2SLS							
Canadian$/US$	1	26.411*	1.066	59.000*	2.549	33.468*	2.465
	2	55.370*	0.797	55.642*	3.045	24.487*	2.761
	3	41.411*	0.888	42.388*	2.923	20.036*	3.095
Mark/US$	1	82.962*	1.084	26.080*	1.778	11.237	3.426
	2	78.926*	1.072	28.664*	1.834	9.117	3.490
	3	78.399*	0.941	42.835*	1.576	11.539	3.114
Yen/US$	1	73.561*	0.370	47.703*	0.817	15.064	1.235
	2	77.417*	0.382	44.767*	0.801	15.888	1.140
	3	74.157*	0.395	37.538*	0.830	17.701*	1.249
Cointegration test results (Sample 2: June 1987–August 1993)							
Differences							
Canadian$/US$	1	51.436*	1.558	46.821*	1.714	29.688*	0.150
	2	58.117*	1.654	44.921*	1.544	28.258*	0.151
	3	58.464*	1.669	45.231*	1.422	26.982*	0.105
Mark/US$	1	22.357*	3.560	30.453*	2.615	7.411	4.148
	2	28.868*	3.758	33.023*	2.707	8.912	4.382
	3	22.151*	2.074	33.856*	2.740	8.518	4.505
Yen/US$	1	51.020*	0.755	—	—	11.538	0.774
	2	51.728*	0.750	—	—	11.755	0.717
	3	51.534*	0.739	20.552*	0.094	11.109	0.814

Table 2 (*Continued*)

	Model	Forecasting horizons					
		1-month ahead		6-month ahead		12-month ahead	
		$r = 0$	$r = 1$	$r = 0$	$r = 1$	$r = 0$	$r = 1$
Error correction							
Canadian$/US$	1	64.098*	1.307	16.278	2.282	27.103*	0.220
	2	56.766*	1.311	42.917*	1.291	18.678	0.105
	3	50.561*	1.330	41.420*	1.331	19.224*	0.202
Mark/US$	1	30.048*	4.242	35.002*	3.172	8.846	2.727
	2	22.756*	7.607	33.280*	3.352	10.938	3.094
	3	23.395*	7.756	35.319*	3.081	9.647	3.306
Yen/US$	1	72.128*	0.596	28.950*	0.005	5.054	0.782
	2	72.884*	0.606	30.592*	0.000	6.216	0.981
	3	72.151*	0.595	30.111*	0.002	6.982	1.055
2SLS							
Canadian$/US$	1	37.035*	1.332	33.724*	0.209	18.697	0.001
	2	48.980*	1.715	28.921*	0.178	12.697	0.018
	3	19.485*	1.611	27.006*	0.112	10.614	0.418
Mark/US$	1	48.015*	6.073	17.852*	3.512	9.603	5.669
	2	43.252*	6.187	18.546*	5.578	10.705	3.137
	3	39.643*	5.870	15.368*	5.636	8.529	2.790
Yen/US$	1	—	—	—	—	—	—
	2	—	—	18.870*	0.248	—	—
	3	—	—	—	—	—	—

Notes: See notes of Table 1. Maximal eigenvalue test statistics for Johansen regressions (lag lengths selected by AIC), where an entry under $r = 0$ indicates the test statistic for the null of $r = 0$ against the alternative of $r = 1$, and an entry under $r = 1$ indicates the test statistic for the null of $r = 1$ against the alternative of $r = 2$. *, indicates significance at 5% MSL using Cheung and Lai (1993a) finite sample critical values. '–' indicates failure to find the same degree of integration between forecast and actual series. See Appendix 2, Cheung and Chinn (1997) for detailed regression results and results based on the trace statistic.

cointegration test only to those series that shared the same order of integration. The results based on the trace statistic are qualitatively similar, and are reported in Appendix 2 of Cheung and Chinn (1997).

For the post-1982 sample at the 1-month ahead horizon, all forecasts are cointegrated with the actual series, except the Canadian dollar/US dollar error correction specification for model 2. At the 6-month ahead horizon, all but two pairs are cointegrated — OLS Model 3 for the mark/dollar rate and the error correction specification of Model 1 for the Canadian dollar/US dollar rate. For the 1-year ahead forecasts, a majority of the pairs fail to reject the null of no cointegration. Interestingly, all the 1-year ahead Canadian dollar/US dollar forecasts are cointegrated with the actual exchange rate.

For the shorter post-Louvre sample involving 1-month ahead forecasts, we find all the pairs (for which both series are I(1)) appear cointegrated. For 6-month ahead forecasts, however, the null of no cointegration is not rejected for one

Canadian dollar/US dollar exchange rate forecast. Moving to the 1-year horizon, a large number of series do not reject the no cointegration null — 19 out of 24 cases for which both series of the pair are I(1). The five series which appear to be cointegrated are once again highly currency specific — in this case, to the Canadian dollar/US dollar rate.

Overall, as the forecast horizon extends out to 12 months ahead, the proportion of cointegrated pairs usually drops drastically: 10 out of 27 in the post-1982 sample. This pattern holds with even greater force for the post-Louvre sample, with only five out of 24 fulfilling the requirement of cointegration.

These 12-month ahead results seem to be specific to currencies. Yen/dollar and mark/dollar pairs seldom appear cointegrated. In fact, most of the cointegrated pairs are Canadian dollar/US dollar. A somewhat disappointing result is that error correction models do not appear to be distinguishable from other specifications in terms of their cointegration characteristics. However, the 1-year ahead horizon is considerably shorter than the 3-year horizon for which Chinn and Meese (1995) found positive results. Indeed, in their study the ECMs did not systematically outperform other estimation methods for shorter horizons.

4.3. Elasticity of expectations

A requirement of forecast consistency is that not only do the forecast and actual series share the same stochastic trend, but also that the cointegrating vector be $(1 -1)$. The results of implementing this test are reported in Table 3. Using the likelihood ratio test on the data from the post-1982 sample, at the 1-month horizon, most of the rejections of unitary elasticity come from forecasts derived from error correction models — seven out of the eight cases reject. The other six are distributed evenly over the OLS and 2SLS specifications. At the 6-month horizon, this pattern is repeated, with six out of eight error correction specifications rejecting unitary elasticities. The other two rejections are for 2SLS specifications. At the 1-year horizon, only one out of the 10 cases rejects — a 2SLS specification of Model 1 for the Canadian dollar/US dollar rate.

Thus, at the 1-month ahead horizon, this restriction is rejected in one half of the cases at the 5% level. At the 6-month ahead horizon, only one-third reject. At the 1-year horizon, only one out of 10 series rejects. However, it is important to note that the number of cointegrated pairs at this horizon is substantially smaller than before. Hence, as the forecast horizon extends forward, the number of cointegrated pairs declines, but of those that are cointegrated, more pass the test of coefficient restrictions.

In the post-Louvre sample, the restriction on the cointegrating vector is only rejected three times, at the 1-month-ahead horizon. This outcome seems to reflect the lower power of the tests given the shorter span of data.

4.4. Discussion

It is important to note how the methodology adopted in this article fits into the

Table 3
Cointegration vector restrictions test

	Model	Sample 1 1983.01–1993.08			Sample 2 1987.06–1993.08		
		1-month ahead	6-month ahead	12-month ahead	1-month ahead	6-month ahead	12-month ahead
Differences							
Canadian$/US$	1	7.61*	1.19	1.21	0.92	0.11	0.60
	2	3.91*	3.22	1.07	0.00	0.19	0.64
	3	2.46	3.57	1.13	0.05	0.16	0.82
Mark/US$	1	4.64*	0.52	—	1.60	0.11	—
	2	1.02	0.17	—	1.43	0.34	—
	3	1.72	—	—	0.55	0.10	—
Yen/US$	1	0.04	0.06	—	2.68	—	—
	2	0.06	0.09	—	2.54	—	—
	3	0.09	0.32	—	2.70	1.36	—
Error correction							
Canadian$/US$	1	8.27*	—	3.24	0.95	—	1.55
	2	—	12.36*	2.69	0.56	0.72	—
	3	11.54*	11.43*	2.46	0.74	1.18	1.19
Mark/US$	1	9.44*	3.64	—	2.18	0.27	—
	2	21.16*	14.27*	—	0.02	0.00	—
	3	14.71*	10.29*	—	0.43	0.40	—
Yen/US$	1	7.81*	7.19*	—	3.92*	0.14	—
	2	2.30	2.99	—	4.51*	0.00	—
	3	4.56*	4.58*	—	3.95*	0.01	—
2SLS							
Canadian$/US$	1	7.05*	14.89*	7.00*	1.51	0.07	—
	2	6.42*	7.82*	2.76	0.36	0.00	—
	3	0.02	0.01	0.09	0.00	0.54	—
Mark/US$	1	2.45	1.14	—	1.41	0.09	—
	2	5.18*	1.44	—	0.76	1.11	—
	3	1.99	1.33	—	0.05	0.15	—
Yen/US$	1	0.09	0.35	—	—	—	—
	2	0.45	0.98	—	—	0.00	—
	3	0.20	0.51	0.66	—	—	—

Notes: See notes of Table 1. The entries are the Likelihood Ratio test statistics for the restriction on the cointegrating vector of $(-1\ 1)$, which is distributed χ^2. A * indicates rejection at the 5% MSL. '—' indicates failure to find the same degree of integration between forecast and actual series, or a failure to find cointegration using the 5% MSL. See Appendix 3, Cheung and Chinn (1997) for detailed results.

extant literature on *ex post* exchange rate forecasting, which uses the random walk as a benchmark. The random walk model will fulfill all three of the consistency criteria set forth, so implicitly it remains the benchmark forecast against which the structural models are compared.

Our results show that it is fairly easy for the generated forecasts to pass the requirement of the same order of integration. The failure of the forecast and the exchange rate to have the same order of integration only accounts for 6% of the rejections. Most of the rejections are attributed to the absence of cointegrating relationship and the non-unitary elasticity of forecasts.

Approximately 26% of the I(1) pairs of forecasts and exchange rates are found to be not cointegrated. Cointegration fails to hold the farther out the forecasts extend. At the 12-month ahead horizon, most exchange rate series and their respective forecasts do not appear cointegrated. The observed pattern does not appear to be completely explained by the decrease in sample sizes and the consequence drop in the power. For the post-1982 sample, the sample size decreases from 123 to 117 (for the 6-month ahead and 12-month ahead forecasts, respectively). On the other hand the rejection rate of the no-cointegration null drops from 25/27 to 10/27. In the case of post-Louvre sample, the observed rejection frequency declines to 5/24 from 22/23, as the number of observations shrinks to 63 from 69.

One possible explanation for this finding is that even though the actual and forecast series are cointegrated, the cointegrating error is so highly autocorrelated, or has such a large variance, that the two series do not appear to be cointegrated. Consistent with the observed pattern of results, the variance of the cointegrating error very likely increases with the forecast horizon, as noted by Clements and Hendry (1994). Another way to interpret this statement is that the Johansen procedure has low power against alternatives where the cointegrating error contains a near unit root. The absence of cointegration between forecasts and spot exchange rates indicates (some of) the exchange rate models considered in our studies do not capture some salient features of exchange rate movements.

Among the cointegrated cases, 22% of them fail the unitary elasticity of forecasts condition. Specifically, the non-unitary elasticity results are found mostly among the 1-month ahead forecasts and those from the error correction specification in the post-1982 samples. Table 4 summarizes these results. In sum, 87 out of the total 162 cases satisfy the consistency requirement.

The pattern of consistency results appears to be currency specific. The Canadian dollar/US dollar forecasts exhibit the strongest evidence of forecast consistency. 36 of 87 consistent forecast series are from Canadian dollar/US dollar exchange rate models. Compared with the yen/dollar and mark/dollar, it may be easier to explain Canadian dollar/US dollar exchange rate movements because of the close linkages, both economic and geographic, between the US and Canada.

Regarding the estimation methodology, the error correction approach generates the least number of consistent forecast series. It accounts for 25% of the consistent cases. This seems to be at variance with results reported in Chinn and Meese (1995). However, it is noted that the horizon considered by them is 3 years while the longest horizon considered in the current study is 1 year.

The choice of model specifications shows no distinguishable effect on the forecast consistency. Of the 87 consistent forecast series, 26 are generated from the flexible price monetary model, 29 from the sticky price model and the remaining are from the model that incorporates the relative price of tradables and non-trada-

Table 4
Summary: consistent forecasts

	Model	Sample 1 January 1983–August 1993			Sample 2 June 1987–August 1993		
		1-month ahead	6-month ahead	12-month ahead	1-month ahead	6-month ahead	12-month ahead
Differences							
Canadian$/US$	1	—	C	C	C	C	C
	2	—	C	C	C	C	C
	3	C	C	C	C	C	C
Mark/US$	1	—	C	—	C	C	—
	2	C	C	—	C	C	—
	3	C	—	—	C	C	—
Yen/US$	1	C	C	—	C	—	—
	2	C	C	—	C	—	—
	3	C	C	—	C	C	—
Error correction							
Canadian$/US$	1	—	—	C	C	—	C
	2	—	—	C	C	C	—
	3	—	—	C	C	C	C
Mark/US$	1	—	C	—	C	C	—
	2	—	—	—	C	C	—
	3	—	—	—	C	C	—
Yen/US$	1	—	—	—	—	C	—
	2	C	C	—	—	C	—
	3	—	—	—	—	C	—
2SLS							
Canadian$/US$	1	—	—	—	C	C	—
	2	—	—	C	C	C	—
	3	C	C	C	C	C	—
Mark/US$	1	C	C	—	C	C	—
	2	—	C	—	C	C	—
	3	C	C	—	C	C	—
Yen/US$	1	C	C	C	—	—	—
	2	C	C	—	—	C	—
	3	C	C	C	—	—	—

Notes: see notes of Table 1. ‘C’ indicates forecasts that pass all three requirements for consistency.

bles. This pattern indicates that the inclusion of additional fundamental variables in the exchange rate equation does not detectably improve forecasting performance at these horizons, a result that corroborates the existing consensus regarding the difficulty in forecasting exchange rates.

These last three observations regarding currency specificity, econometric and economic specifications also hold true for any given forecast horizon. Hence, one can conclude that the numerical tallies are not being driven by particular results that obtain at only the shortest, or longest, horizons.

5. Horvath–Watson test results

In the previous portion of the article, we have adopted a sequential testing procedure, wherein the testing for cointegration and then testing a specific cointegrating vector, are conducted separately. An alternative procedure is to collapse these two steps into one. The Horvath and Watson (1995) methodology is well suited to this task, since it tests the null hypothesis of no cointegration against the alternative of (in this case) cointegration with a vector of *known* coefficients.

Table 5
Horvath–Watson Test Results

	Model	Sample 1 January 1983–August 1993			Sample 2 June 1987–August 1993		
		1-month ahead	6-month ahead	12-month ahead	1-month ahead	6-month ahead	12-month ahead
Differences							
Canadian$/US$	1	—	C	C	C	C	C
	2	—	C	C	C	C	C
	3	—	C	C	C	C	C
Mark/US$	1	—	—	—	—	—	—
	2	—	—	—	C	—	C
	3	—	—	—	C	—	C
Yen/US$	1	—	—	—	—	—	C
	2	—	—	—	—	—	C
	3	—	—	—	—	—	C
Error correction							
Canadian$/US$	1	—	—	C	—	C	C
	2	C	C	C	—	—	C
	3	—	—	C	—	—	C
Mark/US$	1	—	—	—	—	—	—
	2	—	—	—	—	—	C
	3	—	—	—	—	—	C
Yen/US$	1	C	—	—	—	—	—
	2	—	—	—	—	—	—
	3	C	—	—	—	—	—
2SLS							
Canadian$/US$	1	C	C	C	C	C	C
	2	—	—	—	—	—	—
	3	—	—	—	—	—	C
Mark/US$	1	—	—	—	—	—	—
	2	—	—	—	—	—	—
	3	—	—	—	—	—	—
Yen/US$	1	—	—	—	—	—	—
	2	—	—	—	—	—	—
	3	—	—	—	—	—	—

Notes: see notes of Table 1 and Table 4. See Appendix 4, Cheung and Chinn (1997) for detailed results.

Essentially, this procedure reduces to applying a Wald test for zero restrictions on the α coefficients in equation (6). In order to select the optimal VAR lag length, we use the AIC. Horvath and Watson (1995) report appropriate critical values for this Wald test in their Table 1.

The consistency results of the Horvath–Watson procedure are reported in Table 5 (specific Wald and AIC statistics and the corresponding selected lag lengths are reported in Appendix 4 of Cheung and Chinn (1997)). The most striking aspect of the table is that there are many fewer cases of consistency: 42 out of 162, vs. the 87 out of 162 indicated by the Johansen procedure. Another feature of the results is that a higher proportion of the identified cases of consistency are at the longer horizons: 52% of the cases of consistency are at the 12-month horizon, while using the sequential method results in only 17% of consistent cases at this horizon.

Since the Horvath–Watson procedure is widely perceived as more powerful than the Johansen procedure, it is surprising that we obtain these results. We make the following comments. First, one should note that the Johansen procedure tests the null of no cointegration against the alternative of cointegration, with some cointegrating vector that is estimated. Then, the likelihood ratio test is applied to the identified cointegrating vector, where the null hypothesis is (1−1) coefficients and the alternative is a cointegrating vector with differing coefficients. In contrast, the Horvath–Watson procedure tests the null of no cointegration *against an alternative of cointegration with a specific cointegrating vector*. Hence, the Horvath–Watson procedure is indeed more powerful against the null provided one has strong priors on the cointegrating vector. As mentioned in the introduction, for a variety of reasons, including measurement error in the variables used in generating forecasts, there is ample reason to believe that these priors are inappropriate. Hence, the choice of the method depends upon how informative one believes the macroeconomic data used to generate the forecasts are.

Once one makes this realization, it is not so surprising that one finds fewer instances of consistency using the Horvath–Watson procedure — in one case a sequential procedure with two differing sets of null hypotheses and corresponding alternative hypotheses is applied and in the other a single-step procedure is applied. The null and the alternative in the latter do not correspond to that found in the former.

6. Concluding remarks

In this study, we have applied a test of rationality looser than that imposed by the typical rational expectations methodology. Specifically, our definition of consistency requires only that the forecast and the actual series be cointegrated (and hence necessarily of the same order of integration), with cointegrating vector (1−1). These criteria are more appropriate for evaluating forecasts generated from structural models which incorporate macroeconomic data. Such macroeconomic data usually impart serial correlation to the forecast series, which invalidates at least one of the standard criteria for rationality.

Forecasts evaluated are 1-month, 6-month and 12-month ahead forecasts for Canadian dollar/US dollar, mark/dollar and yen/dollar rates. These exchange rate forecasts are generated from three commonly used structural exchange rate models. Three different estimating methods and two forecasting periods are considered.

We find that it is fairly easy for the generated forecasts to pass the first requirement of consistency that the series be of the same order of integration. However, using the Johansen procedure cointegration fails to hold the farther out the forecasts extend. At the 12-month ahead horizon, most series and their respective forecasts do not appear cointegrated. Of the cointegrated pairs, the 1-month ahead forecasts and those from the error correction estimating method tend to reject the restriction of unitary elasticity of forecasts with respect to actual. Overall, 87 out of 162 cases satisfy the requirement of consistency. In terms of the model performance, our results show that approximately half of the forecasts generated by each of the three structural models are consistent; that is they have a one-to-one relationship with the actual exchange rates in the long-run.

Using a Horvath–Watson procedure which imposes a unitary coefficient restriction, we find fewer instances of consistency (42 vs. 87), but a relatively higher proportion of the identified cases of consistency are found at the longer horizons (52% vs. 17%). Although we have forwarded reasons for some of these results, there is certainly call for further investigation. Obviously, neither of these sets of results constitute ideal performance. However, the results indicate these structural exchange rate models are capable of generating forecasts that are related to the actual series in the long-run.

It would be interesting to investigate further why it is so difficult for such forecasts to pass the weak conditions that comprise our concept of consistency. Some plausible candidates include time varying parameters and structural breaks. Further insights into our empirical results may be related to the treatment of trends in the cointegration analysis. However, assessment of these potential explanations is beyond the scope of this article, and is reserved for future research.

Acknowledgements

Useful comments were received from Frank Diebold, Michael Horvath, Kon Lai, John Rogers, participants at the 1996 Winter Econometric Society meetings, the co-editor Michael Melvin and two anonymous referees. We thank Tuan Tran for research assistance and the UC Pacific Rim Research Program for financial support. This article is an abbreviated version of *NBER Working Paper* no. 5943. All remaining errors are solely the authors' responsibility.

Appendix 1: Data

In general, the data are seasonally unadjusted monthly data, derived from

OECD *Main Economic Indicators* (*MEI*). The data sample covers the period June 1973 to August 1993.

Variable	Units	Notes
Exchange rates	US$/foreign currency unit	end-of-period
Narrow money	M1, in billions	end-of-period
Industrial production	1985 = 100	all manufacturing production
Interest rate	in decimal form	3-month CD rate (Can., US); Frankfurt rate (Ger.); call money rate (Japan).
Consumer price index (CPI)	1985 = 100	
Producer price Index (PPI)	1985 = 100	
π	log(CPI) − log[CPI(−4)]	
q	log[(PPI/CPI)/(PPI*/CPI*)]	

References

Cheung, Y.-W., 1993. Long memory in foreign-exchange rates. J. Bus. Econ. Stat. 11, 93–101.

Cheung, Y.-W., Chinn, M.D., 1997. Integration, cointegration and the forecast consistency of structural exchange rate models. NBER Working Paper no. 5943.

Cheung, Y.-W., Lai, K.S., 1995. Lag order and critical values of the augmented Dickey–Fuller test. J. Bus. Econ. Stat. 13, 277–280.

Cheung, Y.-W., Lai, K.S., 1993a. Finite-sample sizes of Johansen's likelihood ratio tests for cointegration. Oxford Bull. Econ. Stat. 55, 313–328.

Cheung, Y.-W., Lai, K.S., 1993b. Long-run purchasing power parity during the recent float. J. Int. Econ. 34, 181–192.

Chinn, M.D., Meese, R.A., 1995. Banking on currency forecasts: how predictable is change in money?. J. Int. Econ. 38, 161–178.

Clements, M.P., Hendry, D.F., 1994. Towards a theory of economic forecasting. In: Hargreaves, C.P. (Ed.), Non-stationary Time Series Analysis and Cointegration. Oxford University Press, Oxford, pp. 9–52.

Diebold, F.X., Nason, J., 1990. Non-linear exchange rate prediction?. J. Int. Econ. 28, 315–332.

Dornbusch, R., 1976. The theory of flexible exchange rate regimes and macroeconomic policy. Scand. J. Econ. 78, 255–275.

Fischer, A.M., 1989. Unit roots and survey data. Oxford Bull. Econ. Stat. 51, 451–463.

Frankel, J.A., 1979. On the mark: a theory of floating exchange rates based on real interest differentials. Am. Econ. Rev. 69, 610–622.

Frankel, J.A., Rose, A.K., 1995. Empirical research on nominal exchange rates. In: Grossman, G., Rogoff, K. (Eds.), The Handbook of International Economics. North-Holland, Amsterdam, Vol. 3, pp. 1690–1729.

Frenkel, J., 1976. A monetary approach to the exchange rate: doctrinal aspects and empirical evidence. Scand. J. Econ. 78, 200–224.

Froot, K., Ito, T., 1989. On the consistency of short-run and long-run exchange rate expectations. J. Int. Mon. Fin. 8, 487–510.

Horvath, M., Watson, M., 1995. Testing for cointegration when some of the cointegrating vectors are prespecified. Econometr. Theory 11, 984–1014.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 58, 1551–1580.

Johansen, S., Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration — with applications to the demand for money. Oxford Bull. Econ. Stat. 52, 169–210.
Liu, P., Maddala, G.S., 1992. Rationality of survey data and tests for market efficiency in the foreign exchange markets. J. Int. Mon. Fin. 11, 366–381.
Lucas, R.E., 1982. Interest rates and currency prices in a two-country world. J. Monet. Econ. 10, 335–360.
Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: do they fit out of sample?. J. Int. Econ. 14, 3–24.
Meese, R.A., Rose, A.K., 1991. An empirical assessment of non-linearities in models of exchange rate determination. Rev. Econ. Stud. 58, 603–619.