

## What Do We Know about Recent Exchange Rate Models? In-Sample Fit and Out-of-Sample Performance Evaluated

Yin-Wong Cheung, Menzie D. Chinn, and Antonio Garcia Pascual

In contrast to the intellectual ferment that followed the collapse of the Bretton Woods era, the 1990s were marked by a relative paucity of new *empirical* models of exchange rates. The sticky-price monetary model of Dornbusch and Frankel remained the workhorse of policy-oriented analyses of exchange rate fluctuations among the developed economies. However, while no completely new models were developed, several approaches gained increased prominence. Some of these approaches were inspired by new empirical findings, such as the correlation between net foreign asset positions and real exchange rates. Others, such as those based on productivity differences, were grounded in an older theoretical literature but given new respectability by the new international macroeconomics (Obstfeld and Rogoff 1996) literature. None of the empirical models, however, were subjected to rigorous examination of the sort that Frankel (1979) and Meese and Rogoff (1983a, b) conducted in their seminal works.

Consequently, instead of re-examining the usual suspects—the flexible price monetary model, purchasing power parity, and the interest differential<sup>1</sup>—we vary the set of performance criteria and expand the set to include the mean squared error, and the direction-of-change statistic. The later dimension is potentially more important from a market timing perspective, besides serving as another indicator of forecast attributes.

To summarize, in this study, we compare exchange rate models along several dimensions:

- Four models are compared against the random walk. Only one of the structural models—the benchmark sticky-price monetary model of Dornbusch and Frankel—has been the subject of previous systematic

analyses. The other models include one incorporating productivity differentials in a fashion consistent with a Balassa-Samuelson formulation, an interest rate parity specification, and a representative behavioral equilibrium exchange rate model.

- The behavior of US dollar-based exchange rates of the Canadian dollar, British pound, German mark, Swiss franc, and Japanese yen are examined. We also examine the corresponding yen-based rates to ensure that our conclusions are not driven by dollar specific results.
- The models are estimated in two ways: in first-difference and error correction specifications.
- In sample fit is assessed in terms of how well the coefficient estimates conform to theoretical priors.
- Forecasting performance is evaluated at several horizons (1-, 4- and 20-quarter horizons), for a recent period not previously examined (post-1992).
- We augment the conventional metrics with a direction-of-change statistic and the “consistency” criterion of Cheung and Chinn (1998).

In accordance with previous studies, we find that no model consistently outperforms a random walk according to the mean squared error criterion at short horizons. However, at the longest horizon we find that the proportion of times the structural models incorporating long-run relationships outperform a random walk is more than would be expected if the outcomes were merely random. Using a 10 percent significance level, a random walk is outperformed 17 percent of the time along a MSE dimension and 27 percent along a direction of change dimension.

In terms of the “consistency” test of Cheung and Chinn (1998), we obtain slightly less positive results. The actual and forecasted rates are cointegrated more often than would occur by chance for all the models. While in many of these cases of cointegration, the condition of unitary elasticity of expectations is rejected; only about 5 percent fulfill all the conditions of the consistency criteria.

We conclude that the question of exchange rate predictability remains unresolved. In particular, while the oft-used mean squared error criterion provides a dismal perspective, criteria other than the conventional ones suggest that structural exchange rate models have some usefulness. Furthermore, structural models incorporating restrictions at long horizons tend to outperform random walk specifications.

## 8.1 Theoretical Models

The universe of empirical models that have been examined over the floating rate period is enormous. Consequently any evaluation of these models must necessarily be selective. The models we have selected are prominent in the economic and policy literature, and readily implementable and replicable. To our knowledge, with the exception of the sticky-price model, they have also not previously been evaluated in a systematic fashion. We use the random walk model as our benchmark naive model, in line with previous work, but we also select one model—the Dornbusch (1976) and Frankel (1979) model—as a representative of the 1970s vintage models. The sticky-price monetary model can be expressed as follows:

$$s_t = \beta_0 + \beta_1 \hat{m}_t + \beta_2 \hat{y}_t + \beta_3 \hat{i}_t + \beta_4 \hat{\pi}_t + u_t, \quad (1)$$

where  $s$  is exchange rate in log,  $m$  is log money,  $y$  is log real GDP,  $i$  and  $\pi$  are the interest and inflation rate, respectively, the caret ( $\hat{\cdot}$ ) denotes the intercountry difference, and  $u_t$  is an error term.

The characteristics of this model are well known, so we will not devote time to discuss the theory behind the equation. We will observe, however, that the list of variables included in (1) encompasses those employed in the flexible price version of the monetary model, as well as the micro-based general equilibrium models of Stockman (1980) and Lucas (1982).

Second, we assess models that are in the Balassa-Samuelson vein, in that they accord a central role to productivity differentials in explaining movements in real, and hence also nominal, exchange rates (see Chinn 1997). Such models drop the purchasing power parity assumption for broad price indexes and allow the real exchange rate to depend on the relative price of nontradables, itself a function of productivity ( $z$ ) differentials. A generic productivity differential exchange rate equation is

$$s_t = \beta_0 + \beta_1 \hat{m}_t + \beta_2 \hat{y}_t + \beta_3 \hat{i}_t + \beta_5 \hat{z}_t + u_t. \quad (2)$$

The third set of models we examine we term the “behavioral equilibrium exchange rate” (BEER) approach. We investigate this model as a proxy for a diverse set of models that incorporate a number of familiar relationships. A typical specification is

$$s_t = \beta_0 + \hat{p}_t + \beta_6 \hat{\omega}_t + \beta_7 \hat{r}_t + \beta_8 \hat{g}debt_t + \beta_9 tot_t + \beta_{10} nfa_t + u_t, \quad (3)$$

where  $p$  is the log price level (CPI),  $\omega$  is the relative price of non-tradables,  $r$  is the real interest rate,  $gdebt$  is the government debt to GDP ratio,  $tot$  is the log terms of trade, and  $nfa$  is the net foreign asset ratio. A unitary coefficient is imposed on  $\hat{p}_t$ . This specification can be thought of as incorporating the Balassa-Samuelson effect, the real interest differential model, an exchange risk premium associated with government debt stocks, and additional portfolio balance effects arising from the net foreign asset position of the economy.<sup>2</sup> Evaluation of this model can shed light on a number of very closely related approaches, including the macroeconomic framework of the IMF (Isard et al. 2001) and Stein's NATREX (Stein 1999). The empirical determinants in both approaches overlap with those of the specification in equation (3).

Models based on this framework have been the predominant approach to determining the level at which currencies will gravitate to over some intermediate horizon, especially in the context of policy issues. For instance, the behavioral equilibrium exchange rate approach is the model that is most used to determine the long-term value of the euro.

The final specification assessed is not a model per se; rather it is an arbitrage relationship—uncovered interest rate parity:

$$s_{t+k} - s_t = \hat{i}_{t,k}, \quad (4)$$

where  $\hat{i}_{t,k}$  is the interest rate of maturity  $k$ . Unlike the other specifications, this relation does not need to be estimated in order to generate predictions.

Interest rate parity at long horizons has recently gathered empirical support (Alexius 2001; Chinn and Meredith 2002), in contrast to the disappointing results at the shorter horizons. MacDonald and Nagayasu (2000) have also demonstrated that long-run interest rates can predict exchange rate levels. On the basis of these findings, we anticipate that this specification will perform better at the longer horizons than at the shorter.<sup>3</sup>

## 8.2 Data and Full-Sample Estimation

### 8.2.1 Data

The analysis uses quarterly data for the United States, Canada, the United Kingdom, Japan, Germany, and Switzerland over the 1973:2 to

