THE FEDERAL RESERVE BANK of ST. LOUIS

WORKING PAPER SERIES

How well do monetary fundamentals forecast exchange rates?

Christopher J. Neely and Lucio Sarno

How well do monetary fundamentals forecast exchange rates?

Christopher J. Neely

Lucio Sarno

May 15, 2002

Abstract: For many years after the seminal work of the Meese and Rogoff (1983a), conventional wisdom held that exchange rates could not be forecast from monetary fundamentals. Monetary models of exchange rate determination were generally unable to beat even a naïve no-change model in out-of-sample forecasting. More recently, the use of sophisticated econometric techniques, panel data, and long spans of data has convinced some researchers (Mark and Sul, 2001) that monetary models can forecast a small, but statistically significant part of the variation in exchange rates. Others remain sceptical, however (Rapach and Wohar, 2001b; Faust, Rogers, and Wright, 2001). It remains a puzzle why even the most supportive studies find such a small predictable component to exchange rates. This article reviews the literature on forecasting exchange rates with monetary fundamentals and speculates as to why it remains so difficult.

INTRODUCTION

In the last decade or so exchange rate economics has seen a number of important developments, with substantial contributions to both the theory and the empirics of exchange rate determination. Important developments in econometrics and the increasing availability of high-quality data have also stimulated a large amount of empirical work on exchange rates. While this research has improved our understanding of exchange rates, a number of challenges and questions remain. One of the most widely studied and still unanswered questions in this literature involves why monetary models of exchange rate determination cannot forecast much of the variation in exchange rates.

The monetary approach to exchange rate determination emerged as the dominant exchange rate model at the outset of the recent float in the early 1970s and remains an important exchange rate paradigm (Frenkel, 1976; Mussa, 1976, 1979; Bilson, 1978). However, Meese and Rogoff's (1983a) finding that monetary models' forecasts could not outperform a simple no-change forecast was a devastating critique of standard models and marked a watershed in exchange rate economics. Even with the benefit of twenty years of hindsight, moreover, evidence that monetary models can consistently and significantly outperform a naïve random walk is still elusive (e.g. Mark and Sul, 2001; Rapach and Wohar, 2001a and 2001b; Faust, Rogers, and Wright, 2001).

This article reviews this puzzle and discusses potential explanations for the consistent failure of monetary models to forecast much variation in nominal exchange rates. We present the essential elements of the monetary model in the next section, and then discuss, in Section 3,

the key empirical studies examining the out-of-sample forecasting performance of the monetary model. Section 4 outlines possible explanations of the apparent failure of monetary model predictions and a final section briefly concludes.

THE MONETARY APPROACH TO EXCHANGE RATE DETERMINATION

In this section we describe the main features of the monetary approach to exchange rate determination in its flexible-price formulation (Frenkel, 1976; Mussa, 1976, 1979).[1]

The monetary approach starts from the definition of the exchange rate as the relative price of two monies and attempts to model that relative price in terms of the relative supply of and demand for those monies. In discrete time, monetary equilibria in the domestic and foreign country respectively are given by:

(1) $\qquad m_t = p_t + \kappa y_t - \lambda i_t$

(2) $\qquad m_t^* = p_t^* + \kappa^* y_t^* - \lambda^* i_t^*$

where $m, p, y,$ and i denote the log-levels of the money supply, the price level, income, and the level of the interest rate, respectively; κ and λ are positive constants; asterisks denote foreign variables and parameters. In the monetary model, the real interest rate is exogenous in the long run and determined in world markets, because of the implicit assumption of perfect capital mobility.

Another building block of the monetary model is absolute purchasing power parity (PPP), which holds that goods-market arbitrage will tend to move the exchange rate to equalize prices in

two countries. For example, if U.S. goods are more expensive than Mexican goods, U.S. and Mexican consumers will tend to purchase more goods in Mexico and fewer in the United States. The increased relative demand for Mexican goods will tend to make the peso appreciate with respect to the dollar and equalize the dollar-denominated prices of U.S. and Mexican goods. The monetary model assumes that PPP holds continuously, so that:

(3) $$s_t = p_t - p_t^*$$

where s_t is the log-level of the nominal bilateral exchange rate (the domestic price of the foreign currency).

The domestic money supply determines the domestic price level and hence the exchange rate is determined by relative money supplies. Subtracting equation (2) from equation (1), solving for $(p_t - p_t^*)$, and inserting the result into equation (3) provides the solution for the nominal exchange rate:

(4) $$s_t = (m_t - m_t^*) - (\kappa y_t - \kappa^* y_t^*) + (\lambda i_t - \lambda^* i_t^*),$$

which is the fundamental equation of the monetary model. The model is often simplified by assuming that the income elasticities and interest rate semi-elasticities of money demand are the same for the domestic and foreign country ($\lambda = \lambda^*$ and $\kappa = \kappa^*$) so that equation (4) reduces to:

(5) $$s_t = (m_t - m_t^*) - \kappa(y_t - y_t^*) + \lambda(i_t - i_t^*)$$

According to equation (5), an increase in the domestic money supply relative to the foreign money stock, for example, induces a depreciation of the domestic currency in terms of the foreign currency. In other words, the nominal exchange rate, s_t, increases. Conversely, a boost in

---

[1] For a more comprehensive discussion of the monetary model and other models of exchange rate determination, see Sarno and Taylor (2002, Chapters 4 and 5).

domestic real income, *ceteris paribus*, creates an excess demand for the domestic money stock. In an attempt to increase their real money balances, domestic residents reduce expenditure and prices fall until money market equilibrium is achieved. Via PPP, the fall in domestic prices (with foreign prices constant) implies an appreciation of the domestic currency in terms of the foreign currency (a rise in the value of domestic currency in terms of foreign currency).

The model further assumes that the uncovered interest parity (UIP) condition holds

(6) $$E_t\left(\Delta s_{t+1}\right) = (i_t - i_t^*)$$

where Δ is the first-difference operator, so that $\Delta x_t = x_t - x_{t-1}$ for any x, and $E_t(\Delta s_{t+1})$ denotes the market expectation of the change in the exchange rate.[2] The expected rate of depreciation of the domestic currency, Δs^e_{t+1}, can then be substituted for the nominal interest rate differential, $(i_t - i_t^*)$ in equation (5) to yield:

(7) $$s_t = (m_t - m_t^*) - \kappa(y_t - y_t^*) + \lambda E_t\left(\Delta s_{t+1}\right)$$

Using the identity $E_t\left(\Delta s_{t+1}\right) = E_t\left(s_{t+1}\right) - s_t$, equation (7) may in turn be rewritten as:

(8) $$s_t = (1+\lambda)^{-1}(m_t - m_t^*) - \kappa(1+\lambda)^{-1}(y_t - y_t^*) + \lambda(1+\lambda)^{-1}E_t\left(s_{t+1}\right)$$

By iterating forward in equation (8), the rational expectations solution to (7) may be written:

(9) $$s_t = (1+\lambda)^{-1}\sum_{i=0}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^i E_t\left[(m_{t+i} - m_{t+i}^*) - \kappa(y_{t+i} - y_{t+i}^*)\right]$$

---

[2] Uncovered interest parity (UIP) states that risk-neutral arbitrage will equalize the expected return on a foreign investment—approximately $\left(E_t\left(\Delta s_{t+1}\right) + i_t^*\right)$—and the return on a domestic investment $\left(i_t\right)$.

where $E_t[\cdot]$ denotes the mathematical expectation conditional on the information set available at time t.[3] It is well known from the rational expectations literature, however, that equation (9) represents only one solution to (7) from a potentially infinite set. In general, given the exchange rate determined according to equation (9), say s^E_t, (7) has multiple rational expectations solutions conforming to:

(10) $$s_t = s^E_t + B_t$$

where the rational bubble term B_t satisfies

(11) $$E_t[B_{t+i}] = \lambda^{-1}(1+\lambda)B_t$$

Therefore, s^E_t simply represents the rational expectations solution to the monetary model in the absence of rational bubbles. Rational bubbles represent significant departures from the fundamentals of the model that would not be detected in a specification such as (5). Thus, testing for the presence of bubbles can be interpreted as an important specification test of the model (Meese, 1986).

Assumptions of the Monetary Model

Although the simplicity of the flexible-price monetary model is very attractive, this simplicity requires many assumptions. Open economy macroeconomics is essentially about six aggregate markets: goods, labor, money, foreign exchange, domestic bonds (i.e. non-money assets) and foreign bonds. The monetary model concentrates, however, directly on equilibrium conditions in only one of these markets, the money market. This is implicitly achieved in the

---

[3] Moving from equation (8) to equation (9) requires writing the expression for s_{t+1} in terms of $E_{t+1}(s_{t+2})$ implied by (8), taking expectations, substituting the result for $E_t(s_{t+1})$ in (8), and then repeating the process for $E_{t+2}(s_{t+3})$, $E_{t+3}(s_{t+4})$, etc.

following fashion. By assuming perfect substitutability of domestic and foreign assets, the

domestic and foreign bond markets essentially become a single market, reducing the number of

markets to five. The exchange rate adjusts freely to equilibrate supply and demand in the foreign

exchange market. Perfectly flexible prices and wages likewise equilibrate supply and demand in

the goods and labor markets. Thus, three of the five remaining markets are cleared. Recalling

Walras' law, according to which equilibrium in *n-1* markets of an *n*-market system implies

equilibrium in the *n*-th market, equilibrium of the full system in the model is then determined by

equilibrium conditions for the money market. The flexible-price monetary model is thus,

implicitly, a market-clearing general equilibrium model in which continuous PPP among national

price levels is assumed (Sarno and Taylor, 2002 (Chapter 4)).

Sticky-Price Monetary Models

The very high volatility of real exchange rates during the 1970s float, cast serious doubts

on the assumption of continuous PPP and inspired the development of further classes of models,

including sticky-price monetary models and equilibrium models.[4]

The sticky-price monetary model, due originally to Dornbusch (1976), allows short-term

overshooting of the nominal and real exchange rates above their long-run equilibrium levels. In

---

[4] Equilibrium exchange rate models, due originally to Stockman (1980) and Lucas (1982), analyze the general equilibrium of a two-country model in a representative agent utility maximizing framework with sound microfoundations. Equilibrium models may be viewed as an extension or generalization of the flexible-price monetary model that allows for multiple traded goods and real shocks across countries. These models are not amenable to direct econometric testing or to the formulation of models designed to forecast exchange rates because they are based on utility functions that cannot be directly estimated. (Rather, researchers have sought to test the broad rather than specific implications of this class of models

this model, it is assumed that there are "jump variables" in the system (exchange rates and interest rates) compensating for stickiness in other variables, notably goods prices. Consider the effects of a cut in the nominal domestic money supply. Since goods prices are sticky in the short run, this implies an initial fall in the real money supply and a consequent rise in interest rates to clear the money market. The rise in domestic interest rates then leads to a capital inflow and an appreciation of the nominal exchange rate. Investors are aware that they are artificially forcing up the value of the domestic currency and that they may therefore suffer a foreign exchange loss when the proceeds of their investment are used to repay liabilities in foreign currency. Nevertheless, as long as the expected foreign exchange loss (the expected rate of depreciation of the domestic currency) is less than the known capital market gain (the interest rate differential) risk-neutral investors will continue to borrow abroad to buy domestic assets. A short-run equilibrium is achieved when the expected rate of depreciation is just equal to the interest rate differential, i.e. when UIP holds. Since the domestic currency must be expected to depreciate because of the interest rate differential, the domestic currency must have appreciated beyond its long-run, PPP equilibrium. In the medium-run, however, domestic prices begin to fall in response to the fall in the money supply. This alleviates pressure in the money market (the real money supply rises) and domestic interest rates start to decline. The exchange rate then depreciates slowly towards long-run PPP. Thus, this model can explain the apparent paradox that the exchange rate of currencies for countries with relatively higher interest rates tend to depreciate: the initial interest rate rise induces a sharp exchange rate appreciation, followed by a slow depreciation as prices adjust, which continues until long-run PPP is satisfied.

---

for exchange rate behavior.) Similar reasoning applies, at least at the present stage, to the literature on "new open economy macroeconomics" (see Lane, 2001; Sarno, 2001, and the references therein).

Christopher J. Neely and Lucio Sarno   7   Draft 2 / 5/15/02

Nevertheless, it should be clear that, regardless of whether one assumes that prices are flexible or sticky, the traditional flexible-price monetary model and its sticky-price formulation imply exactly the same fundamental equation for the exchange rate, which is of the form (5). We now turn to the empirical evidence on the performance of the monetary model in forecasting exchange rates.

FORECASTING EXCHANGE RATES WITH MONETARY MODELS

The move to floating exchange rates in the 1970s spawned a wealth of theoretical efforts to explain their observed high volatility. The monetary models discussed in the previous section were among the most popular and intuitively appealing. It was natural to examine the empirical fit and forecasting ability of these models. This section selectively reviews the long literature attempting to use monetary models to forecast exchange rates.[5]

Meese and Rogoff (1983a and 1983b)

Meese and Rogoff (1983a, 1983b)—hereafter MR—conducted the seminal work in the use of monetary models to forecast the exchange rate. Their procedure was straightforward: They regressed the log of exchange rates on various combinations of relative macroeconomic

---

[5] This paper focuses on forecasting exchange rates with monetary models. There are many non-monetary models available, however. Fair (1999) uses a non-monetary macro model; Clarida and Taylor (1997) and Clarida, Sarno, and Taylor (2003) use models based on the term structure; and Evans and Lyons (1999) use order flow models to explain exchange rate changes.

variables typically included in the exchange rate models of the 1970s.[6] The basic prediction

equation was as follows:

$$
\begin{aligned}
(12) \quad s_{t+k} = a_0 &+ a_1\left(m_t - m*_t\right) + a_2\left(y_t - y*_t\right) + a_3\left(i_t - i*_t\right) \\
&+ a_4\left(\pi^e_t - \pi^e*_t\right) + a_5 tb_t + a_6 tb*_t + u_t
\end{aligned}
,
$$

where s_t, m_t, y_t, i_t, π^e_t, and tb_t are the logs of the exchange rate, domestic (U.S.) money supply,

output, interest rates, expected inflation, and the trade balance, at time t. Asterisks denote

foreign variables. MR interpreted exchange rate models, such as the Frenkel-Bilson, Dornbusch-

Frankel and Hooper-Morton models, as implying different sets of restrictions on the coefficients

in the regression (Hooper and Morton, 1982).   As is the case with most estimation of

macroeconomic models, little effort was made to explicitly map the model to the functional form

and estimation procedure.

The data were monthly observations from March 1973 through June 1981. MR estimated

the models on in-sample periods by several techniques, including ordinary least squares (OLS),

generalized least squares (GLS), to correct for serial correlation in the errors, and Fair's (1970)

instrumental variables (IV), to correct for simultaneous equations bias.[7] To allow the out-of-

sample forecast coefficients to change, rolling regressions with fixed sample sizes were used.

That is, coefficients were initially estimated using data until November 1976, then 1-, 3-, and 12-

month forecasts were constructed. To construct the next set of forecasts, the next month of data

(December 1976) was added, the first month of data were dropped and the coefficients were

reestimated. For the exercises in which future values of the independent variables were needed

---

[6] MR also estimated univariate models of exchange rate changes and vector autoregressions, employing all the variables in equation (12). These models were also unsuccessful, however, and this paper focuses on efforts with monetary models.

[7] MR were very thorough in checking the robustness of their results to changes in procedures; because of space constraints, it is difficult to list all the permutations of models, estimation methods, and data in this paper.

to construct forecasts, MR provided the models with actual future values of the independent variables—instead of forecasting them—to give the monetary model the best possible chance of forecasting well.[8]

MR used both in-sample model evaluation criteria, such as the R^2, and out-of-sample criteria, such as the comparison of the root-mean-squared error (RMSE) of the model's forecast with that of a benchmark forecast, the driftless random walk. Many of the estimated models fit the in-sample data well. In-sample evaluation techniques, which permit the use of all the data available to the researcher, provide more precise estimates of statistics of interest and therefore have greater *power* to reject the null hypothesis of no predictability of the exchange rate.[9] The advantage of out-of-sample evaluation procedures is that they implicitly test the stability of the estimated coefficients and therefore provide a more stringent and realistic hurdle for models to overcome.

The main conclusion of the MR paper was that none of the structural exchange rate models were able to forecast out-of-sample better than a naive no-change forecast by mean-squared-error (MSE) and mean-absolute-error (MAE) criteria. There was some evidence of predictability at longer horizons, but—given the massive failure at short horizons—this did not receive much attention.

---

[8] Ironically, Faust, Rogers, and Wright (2001) have recently shown that real-time, Federal Reserve forecasts of future independent money and output variables actually generate better forecasts of the future exchange rate than do actual future values of the independent variables.

[9] The *power function* of a statistical test is the probability of rejecting the null hypothesis, conditional on the true data generating process. The *size* of a test is the maximum power under the null hypothesis.

Econometric Problems

The MR exercise had a number of econometric problems, many of which they recognized and attempted to mitigate with variations on their procedures. First, because the explanatory variables were all endogenous—determined within the economic system—the estimated coefficients in the equations surely suffered from simultaneous equations bias. That is, even with an arbitrarily large amount of data, the coefficient estimates would not converge to any structural parameters. MR (1983b) attempted to correct for this problem with IV estimation and an in-sample grid search over possible parameter values. The IV estimation did not help and an in-sample grid search constituted unconvincing evidence. Because the benchmark no-change prediction is nested within the model, some combination of parameters must perform at least as well as the no-change model within the sample. A model with all zeros for coefficients, for example, will perform exactly as well as the no-change forecast. And, in practice, some combination of coefficients will almost certainly outperform the no-change forecast, in-sample.

It is unclear, however, why biased coefficients would be a problem for a forecasting exercise. If the covariance matrix of the structural errors is homoskedastic and stable over time, forecasts from biased coefficients would be superior to those from structural parameters.

Second, problems with the persistence of the variables in the regression were not dealt with in a convincing way. (See the boxed insert on persistence and cointegration.) In particular, because the dependent variable, the log exchange rate, in the predictive regression was probably

I(1), but not cointegrated with any combination of the independent variables, the error term was probably I(1) and the coefficient estimates were inconsistent, meaningless for forecasting.

More Negative Results

The very strong negative results of the MR study spawned an enormous amount of subsequent research that varied econometric techniques or the information set to try to rescue the ability of fundamental models—or any models—to forecast exchange rates. For example, Wolff (1987, 1988), Schinasi and Swamy (1989), and Canova (1993) used time varying coefficients in an effort to compensate for instability in the model. Other authors used expanded information sets or different functional forms for the empirical work (Meese and Rose, 1990). Such efforts, however, proved immediately unsuccessful or were subsequently shown to be fragile to minor changes in technique or the data. The MR result that monetary fundamentals do not help predict exchange rates remained conventional wisdom.

Resurrecting the Monetary Approach: Mark (1995)

Some progress was made, however, in the 1990s. Mark (1995) and Chinn and Meese (1995) focused on neglected aspects of the problem, primarily the possibility of better long-run predictability, but also the finite sample properties of the test statistics, and the estimation procedure implied by the theoretical model.

Mark (1995) considers an expression relating the change in the exchange rate to the deviation of the exchange rate from a linear combination of relative money and relative output,

which is called the *fundamental* value of the exchange rate. Essentially Mark exploits the

monetary model equation (5), assuming that $\kappa=1$ and the interest differential is equal to zero, so

that the fundamentals term is

(13) $$f_t = \left[(m_t - m^*_t) - (y_t - y^*_t)\right]$$

In this model, the difference between the current fundamentals and the current exchange rate—

called the *error correction term* $(f_t - s_t)$—determines the k-period ahead change in the exchange

rate:

(14) $$s_{t+k} - s_t = \alpha_k + \beta_k (f_t - s_t) + \upsilon_{t+k,t}$$

Using quarterly data on U.S. dollar exchange rates with Canada (CAD), Germany (DEM), Japan

(JPY), and Switzerland (CHF) from 1973:2 to 1991:4, Mark computed the forecast regression in

(14) over horizons of 1, 4, 8, 12, and 16 quarters. Out-of-sample forecasts began in 1984:1.

As did MR, Mark (1995) relied on both in-sample and out-of-sample evidence to assess

the degree of predictability in his model. In-sample evidence included both the statistical

significance of the estimated coefficients, β_k, and the R^2s of the regressions at various horizons,

R_k^2. Out-of-sample evidence focused on the root-mean-squared-error provided by the forecasts

from the estimated model versus those from a driftless random walk, the same out-of-sample

benchmark as used by MR. Mark (1995) examined both the ratio of these RMSEs, OUT/RW$_k$—

also called the Theil U statistic—and the statistical significance of their difference, the DM

statistic (Diebold and Mariano, 1995).[10]

---

[10] Mark (2001) provides an easily readable discussion of exchange rate forecasting issues.

In evaluating the statistical significance of the results, Mark (1995) confronted some of the same econometric problems that beset MR. Two of the complications are well known: 1) Because the independent variable $(f_t - s_t)$ is highly autocorrelated, the coefficients α_k and β_k would be biased in finite samples (Mankiw and Shapiro, 1986; Stambaugh, 1986); and 2) for forecast horizons (k) greater than one period, the overlapping forecast errors would have autocorrelation of at least degree k-1, complicating the construction of functions of the forecast errors, like coefficient standard errors.

To overcome these difficulties, Mark (1995) relied on *nonparametric bootstrapping* to determine the statistical significance of his in-sample and out-of-sample results. Nonparametric bootstrapping is a method of simulating the distribution of statistics with the distribution of actual errors estimated by the model—rather than pseudo-random errors from a normal (or other) distribution—under some assumption about how the data were generated (Berkowitz and Kilian, 2000). In this case, a bootstrapping exercise calculates how often an economy in which there was no predictability would produce as much support for predictability as found in actual data. In other words, the question is: Do the real data look like they might have been generated by a model in which there is no predictability?

Mark (1995) assumed a null data generating process (DGP) where the exchange rate change is simply a constant plus an error term and the error correction term (ECT) $(f_t - s_t)$ follows an autoregressive process of order P,

(15) $$\Delta s_t = a_0 + \varepsilon_{1,t},$$

$$(16) \qquad \left(f_t - s_t\right) = b_0 + \sum_{j=1}^{P} b_j \left(f_{t-j} - s_{t-j}\right) + \varepsilon_{2,t}.$$

Note that (15) implies that the change in the exchange rate is unforecastable. Mark chose the parameters of the data generating process (DGP) to match the actual data and then constructed bootstrapped distributions for the test statistics as follows:

1. Estimate the null DGP, which is described by equations (15) and (16).

2. Draw 2000 error samples of size 76 from the estimated null DGP covariance matrix.

3. Use the errors to compute 2000 series of s_t and f_t-s_t, from equations (15) and (16).

4. Run the predictive regression, equation (14), to obtain estimates of β_k, its t-statistics, and the out-of-sample RMSEs from the estimated models and the benchmark no-change prediction.

Table 1 presents selected results from Mark's (1995) exercise with significance levels generated from a DGP described by (15) and (16). Out-of-sample forecasts were evaluated against the benchmark of a driftless random walk—no change in the exchange rate. Mark (1995) concluded that evidence of predictability, including β_ks, adjusted β_ks, t-statistics, and R^2s, increase with the forecast horizon and that there is evidence of statistically significant forecastability at horizons of 12 and 16 quarters for the Deutsche mark (DEM) and Swiss franc (CHF). In the German case, for example, the β_1 is 0.035 and the β_{16} is 1.324. The t-statistics (p-values) likewise rise (fall) with k, except for Canada. The p-values for the German β_1 and β_{16} t-statistics are 0.291 and 0.038, respectively. Likewise, the strongest out-of-sample evidence for predictability is at the longest horizons. The OUT/RW statistics—which are less than one when the monetary forecasting regression provides lower RMSEs than the no-change forecast—show that the monetary model beats the benchmark at every horizon for the CHF and the JPY and at

the 12- and 16-quarter horizons for the DEM. In the latter case, the regression's RMSE is about

½ that of the no-change forecast at the 16-quarter horizon.

Unpredictability Strikes Back

Mark's (1995) innovative use of the bootstrap solved a number of econometric problems

and appeared to show that there was greater power to predict exchange rates at long horizons

than at short horizons. And, his conclusions were largely buttressed by those of Chinn and

Meese (1995), who investigated many of the same issues and used a wider variety of explanatory

variables, including trade balance, the relative price of tradeables/nontradeables, interest rates,

and inflation, as well as nonparametric methods. Chinn and Meese (1995) found that their

fundamental-based error correction models outperformed the random walk model for long-term

prediction horizons.

Soon, however, other researchers, such as Berkowitz and Giorgianni (2001) and Kilian

(1999) began to criticize Mark's (1995) methodology and the resultant conclusions. Berkowitz

and Giorgianni (2001) focused on how Mark's (1995) assumptions about the long run behavior

of the data series influenced the evidence of predictability. Kilian (1999) looked more carefully

at the form of the assumed data generating process and the robustness of the results to changes in

the sample. Both criticisms focused on a disadvantage of bootstrapping and other simulation

procedures: The results can depend crucially on the assumed DGP.

Berkowitz and Giorgianni (2001)—hereafter BG—pointed out that Mark's (1995) DGP—equations (15) and (16) in this paper—implicitly assumed that the exchange rate and the macroeconomic fundamentals were cointegrated, meaning that while each of the series $\{s_t, f_t\}$ might be individually I(1), a linear combination of them is stationary, or I(0). In other words, even if the difference between f_t and s_t is nonstationary in the real data, estimation of equation (16) will tend to generate data in which the difference between f_t and s_t is stationary. The generated exchange rate, s_t, cannot diverge very far from the generated macroeconomic fundamental, f_t. Ex ante, it is not obvious that cointegration is an important issue, as cointegration is neither a necessary nor sufficient condition for fundamentals to predict exchange rate changes. However, in this case, BG argue that the distribution of the test statistics depends on whether there is cointegration or not.[11] If the f_t and s_t are not cointegrated in the real data, then the critical values generated under the assumption of cointegration will be incorrect. The critical values will be incorrect because the forecasting regression, (14), is almost a *spurious regression* (see the boxed insert on persistence) because, as the forecast horizon, *k*, increases, the change in the exchange rate, Δs_{t+k} in (14), becomes more persistent and—if there is no cointegration between s_t and f_t—the independent variable (f_t-s_t) is I(1). When both sides of the forecasting equation are highly persistent, it approaches a *spurious regression*, in which estimated coefficients falsely appear to be statistically significant. More generally, the distribution of the estimated coefficient from equation (14) will depend on the degree of persistence in the regressor (f_t-s_t). If the null DGP fails to model the persistence of the regressor

---

[11] Berben and van Dijk (1998) derive the asymptotic distributions of the estimator of the regression parameter and its t-statistic, under the null hypothesis of no cointegration. They find that the distribution does not depend on the forecast horizon; long-horizon regressions have no power advantages in testing for cointegration. Their analysis shows that Mark's (1995) results can be at least partly explained by his assumption of cointegration.

(f_t-s_t) correctly, then the critical values of the forecasting statistics will be wrong and the inference drawn from the test might be incorrect.

Table 2 presents the results of a Horvath-Watson test (see the boxed insert on persistence) conducted by BG for cointegration of the exchange rate and macro fundamentals. BG were able to reject the null of no cointegration for only one rate, the CHF. Kilian's results were even more negative toward the cointegration hypothesis; his test failed to reject the null of no cointegration for any exchange rate.[12] Unfortunately, it is often impossible to reject the null of no cointegration, even if (f_t-s_t) is stationary but persistent. To better evaluate the balance of the evidence, Kilian adapts an idea of Rudebusch (1993) to weigh the balance of evidence for and against cointegration.[13] Kilian finds that the evidence favors the cointegration hypothesis for the CHF, was indeterminate for CAD, and favored the null of no cointegration for the DEM and JPY. Nevertheless, Kilian (1999) concludes that the data are potentially consistent with either the assumption of cointegration between s_t and f_t or the contrary, no cointegration. He notes that even if the series are cointegrated, the ECT (f_t – s_t), reverts to its mean very slowly.

Because the Horvath-Watson test results imply that Mark's (1995) assumption of cointegration might be incorrect and because this assumption might influence the distribution of test statistics, BG reexamined the forecastability question without the cointegration assumption. In particular, BG conducted two bootstrapping experiments to study the behavior of the system under alternative assumptions about the DGP. Their first model assumed that the exchange rate

---

[12] Kilian's (1999) Horvath-Watson results might have been different because his sample and estimation methods were different than BG's.
[13] Rudebusch (1993) examines whether one can differentiate the short-run persistence properties under the best stationary model and the best nonstationary model.

is a random walk with drift—as did Mark (1995)—and that macro fundamentals (f_t) follow an AR(3) process.

$$(17) \qquad\qquad \Delta s_t = a_0 + \varepsilon_{1,t},$$

$$(18) \qquad\qquad f_t = b_0 + \sum_{j=1}^{3} b_j f_{t-j} + \varepsilon_{2,t}.$$

This first model did not assume cointegration and, in generating data, the covariance between the error terms $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ was set equal to zero. The exchange rate and macro fundamentals were independent by construction.

Table 3, which is excerpted from Table 4 in BG, shows the results of Mark's forecasting exercise with three changes: 1) P-values were calculated with the DGP described in (17) and (18); 2) Parametric bootstrapping was used in place of nonparametric bootstrapping to generate data; 3) The out-of-sample benchmark included a drift term.[14] BG find that many of the DEM statistics—denoted by shaded boxes—are no longer significant.[15] Only the CHF shows much evidence of predictability.

Table 4—excerpted from Table 5 in BG—is constructed in exactly the same way as Table 3, except that it extends the sample by three years to the end of 1994. With this change, there is now no evidence of predictability at the 5% level, even at long horizons for any exchange rate except the CHF.[16] However, there is now evidence of predictability in the OUT/RW statistics at shorter horizons for the CHF.

---

[14] Kilian (1999) emphasized the importance of a drift in the out-of-sample benchmark, as discussed below.

[15] Tables 3, 4, and 5 show only a subset of the test statistics.

[16] The overturned results were from t(A) and DM(A, 20) statistics, some of which Table 4 does not show.

The second BG model was an unrestricted vector autoregression (VAR) for the pair $\{s_t, f_t\}$. BG considered this model, which permitted but did not enforce cointegration, as an intermediate case between the assumption of cointegration enforced by Mark and the assumption of independence that produced Tables 3 and 4.

$$(19) \qquad s_t = a_0 + \sum_{j=1}^{P} a_{1j} s_{t-j} + \sum_{j=1}^{P} a_{2j} f_{t-j} + \varepsilon_{1,t}$$

$$(20) \qquad f_t = b_0 + \sum_{j=1}^{P} b_{1j} s_{t-j} + \sum_{j=1}^{P} b_{2j} f_{t-j} + \varepsilon_{2,t}.$$

The results from the unrestricted VAR—shown in Table 5—show very little evidence of predictability except for the CHF. BG noted that for the DEM, JPY, and CAD the p-values for the OUT/RW statistics are smaller at shorter horizons than they are at longer horizons, indicating less evidence of predictability at longer horizons, in contrast to Mark's basic conclusion.[17]

Kilian's (1999) primary focus was the study of the *power function* of forecasting regressions at short horizons versus long horizons rather than foreign exchange predictability per se. Such a study of power requires a null DGP. Kilian (1999) carefully mapped the monetary model to a constrained vector error correction model (VECM), which he estimated by feasible generalized least squares (FGLS) to construct bootstrapping distributions for the forecasting statistics.

$$(21) \qquad \Delta s_t = v_e + u_{1,t}$$

---

[17] The unrestricted VAR does permit predictability, so the p-values in Table 5 are the probabilities of obtaining the test statistics, conditional on exchange rates and fundamentals following the estimated VAR.

(22) $$\Delta f_t = v_f - h_2[s_{t-1} - f_{t-1}] + \sum_{j=1}^{p-1} \xi_j^{21} \Delta s_{t-j} + \sum_{j=1}^{p-1} \xi_j^{22} \Delta f_{t-j} + u_{2,t}.$$

where the system requires $h_2 < 0$ for stability. Kilian notes that while it is asymptotically

equivalent to Mark's (1995) approximation, (21) and (22) will generate a different small sample

distribution because of the different lag structure and estimation procedure. The results will also

be sensitive to whether the estimated coefficients for the null DGP have been corrected for the

finite-sample bias (see the boxed insert) caused by the persistence of the regressors, which was

not compensated for in Mark's procedure.

Kilian (1999) also emphasized the importance of the treatment of the drift term in the

forecasting procedures. Specifically, Mark's inconsistency in permitting a drift in the bootstrap

DGP but not in the benchmark forecast biased the bootstrap critical values. In addition, to

compare a fundamental forecast with drift to a driftless random walk conflates the contributions

of the fundamentals and the drift. That is, if the out-of-sample statistics for the fundamental

model are superior to those of the driftless random walk, one cannot be certain whether it was the

contribution of the fundamentals or just the drift in the monetary model. To isolate the marginal

contribution of the fundamentals, one has to allow for a drift in the benchmark forecast.

To illustrate the importance of the treatment of the drift term, Table 6—excerpted from

the December 1997 working paper version of Kilian (1999)—presents the results from Mark's

(1995) forecasting exercise using Kilian's DGP—equations (21) and (22)—with and without a

drift in the random walk. The left-hand panel presents results from the driftless random walk

benchmark while the right-hand panel presents results from the random walk with drift.

Contrasting the results, using a drift in the benchmark eliminates any evidence of predictability

from the JPY case but increases the predictability in the CAD and CHF, especially at short horizons.

Indeed, both BG and Kilian (1999) took issue with the whole idea of finding predictability in long-horizon regressions. BG show that there will be no more predictability at long horizons than at short horizons with a linear model. Kilian (1999) focused on the question of whether long-horizon regressions truly have greater power to find predictability than short-horizon tests. In particular, he extended the analytic work done by BG with Monte Carlo experiments that showed that increasing evidence of predictability at long horizons was due to the fact that such tests were more likely to err in favor of finding predictability where there was none, rather than really being better at finding latent predictability. In econometric jargon, the results were due to *size distortions* rather than *power gains*. To summarize: Both BG and Kilian (1999) conclude that it doesn't help to increase the forecast horizon if the DGP is linear.

The work of BG and Kilian (1999) shows that the results of the forecasting exercise were sensitive to a number of factors, including the data sample, the assumption of cointegration in the DGP, the lag structure of the DGP, the benchmark for out-of-sample comparison, and whether one corrects the DGP for bias generated by persistent regressors.[19] Indeed, their conclusions on predictability are very similar. BG conclude that failure to impose cointegration leaves only weak evidence of predictability and that is at predominantly short horizons. Kilian (1999) similarly concludes that with properly generated critical values, there is some evidence that

---

[19] Groen (1999) also reports the fragility of Mark's (1995) results to the chosen sample.

monetary fundamentals predict foreign exchange rates but no evidence of more forecastability at longer horizons.

Panel Studies

When alternative explanations—i.e., predictability or no predictability—seem to fit the data equally well, employing additional data often illuminates the issue. In the present case, one might combine evidence from many exchange rates in a panel study of predictability, under the assumption that the exchange rates are either predictable from fundamentals for all the countries or predictable for none of them. Groen (2000) and Mark and Sul (2001) aggregated information about the predictability of exchange rates across countries. Rapach and Wohar (2001b) have examined whether such aggregation is appropriate.

Groen (2000) examines the question of whether exchange rates are cointegrated with fundamentals using both rate-by-rate Johansen (1991) cointegration tests and the Levin and Lin (1993) panel unit root tests on 4 subsets of 14 exchange rates against the U.S. dollar (USD): 1) all 14; 2) the G-10; 3) the G-7; and 4) the European Monetary System (EMS). The rate-by-rate Johansen tests reject the null of no cointegration in about 1/3 of the cases using either the USD or the DEM as numeraire currency; this suggests that "cointegration isn't widespread." The more powerful Levin-Lin (1993) panel test, however, rejects the null of no cointegration jointly for all the rates at the 5% level, using either the USD or DEM as numeraire for the 14 country panel.[20]

---

[20] Taylor, Peel, and Sarno (2001) note that panel unit root tests tend to reject the null of a unit root if even one of the series is stationary, because the null hypothesis is that all of the series have a unit root.

While the results for smaller panels are often insignificant, the overall conclusion is that the most powerful tests are supportive of cointegration.

The growing consensus against long-horizon prediction regressions and the econometric complications caused by overlapping forecast errors led Mark and Sul (2001) to eschew the search for long-run predictability in favor of a one-period ahead panel regression of quarterly data on 18 exchange rates and fundamentals. The sample started in 1973:1 and ended in 1997:1. Mark and Sul (2001) first test for and find evidence of cointegration with a panel dynamic OLS framework, which controls for asymptotic bias in the forecast statistics. This cointegration finding is used to construct the bootstrapped data that corrects coefficients for persistent regressor bias and evaluates the statistical significance of Theil U statistics. The forecasting equation is the multi-exchange rate analogue to those used in previous papers, a one-period ahead panel regression, estimated by seemingly unrelated regressions (SUR) over expanding samples.[21] The system could be written as follows:

(23) $$s_{i,t+1} - s_{i,t} = \beta\left(f_{i,t} - s_{i,t}\right) + \varepsilon_{i,t+1}$$

$$\varepsilon_{i,t+1} = \gamma_i + \theta_{t+1} + u_{i,t+1} \quad ,$$

where $s_{i,t}, f_{i,t}$ are the log exchange rate and the log fundamentals of exchange rate i at time t, γ_i is an exchange rate specific error, θ_t is a time specific error and $u_{i,t}$ is an idiosyncratic error. Three exchanges rates are considered as numeraire for the system: the USD, the CHF, and the JPY. In the one-step ahead forecasting exercises, Mark and Sul (2001) find that monetary fundamentals have a small but statistically significant amount of predictability—using Theil U statistics—

---

[21] In expanding samples, one-period is added to the in-sample data used to estimate coefficients before each forecast to give the model the maximum amount of data with which to construct out-of-sample forecasts.

when the USD or CHF is numeraire but none when the JPY is the standard. They also find that monetary fundamentals predict somewhat better than PPP fundamentals.

Both the Groen (2000) and Mark and Sul (2001) studies pooled data across countries to try to bring more power to answering the question of how well monetary models predict the exchange rate? The practice of pooling data across countries assumes, of course, that the same data generating process produces the data for all the countries. Such assumptions are called "homogeneity assumptions." If the data generating process is different across countries, however, then falsely pooling the data can lead to incorrect inference. Using the Mark and Sul (2001) data set, Rapach and Wohar (2001b) first confirm previous results that the monetary model fits very poorly in country-by-country estimation during the floating rate period (1973:1-1997:1). "It is difficult to overstate how poorly the monetary model performs…on a country by country basis" (Rapach and Wohar, 2001b, p. 3). In contrast, however, pooled estimates do support the monetary model, as in Mark and Sul (2001). Next, the authors formally test whether the cross-country homogeneity assumptions are justified. That is, is it likely that one data generating process could have produced the disparate coefficient estimates from the 14 different exchange rates? A Wald test rejects this one-DGP hypothesis for most subsets of countries (Mark, Ogaki, and Sul, 2000). And a Monte Carlo study shows that it is very plausible that heterogeneous data generating processes—fit to the 14 exchange rate/fundamental processes— could produce pooled parameter estimates similar to those found in the real data. These findings cast doubt on the wisdom of pooling data across countries or the reliability of the conclusions.

Arguing in favor of pooling, however, is the fact that the pooled parameter estimates are as good as the country-by-country forecasts at short horizons and better at long horizons. Rapach and Wohar (2001b) cite Pesaran, Shin and Smith (1999) as arguing that omitted variables and measurement error might lead to the false rejection of homogeneity restrictions and that pooling might still be appropriate and helpful under such circumstances. Ultimately, Rapach and Wohar (2001b) conclude that researchers could reasonably differ about the fit of the monetary model of exchange rates during the Post-Bretton Woods period.

Long Spans of Data

Combining evidence from many countries in a panel study is one way to increase the available data to determine whether exchange rates are cointegrated with fundamentals. Another approach is to simply use much longer spans of data. Rapach and Wohar (2001a) took this latter approach, using exchange rate, money, and output data from 14 industrialized countries, over a span as long as 115 years (1880-1995), to investigate the long run relationship between these variables. Table 7 summarizes the results from Rapach and Wohar (2001a).

First, Rapach and Wohar (2001a) noted that if exchange rates are to be predicted from relative money and output, some combination of the three variables $\{s_t, (m_t - m_t^*), (y_t - y_t^*)\}$ must be stationary (I(0)). If no combination is I(0), then the error from any forecast will become arbitrarily big as time goes on, bigger than the benchmark error. If one of these variables $\{s_t, (m_t - m_t^*), (y_t - y_t^*)\}$ is I(1), for example, while the other 2 are I(0), then no linear combination can be I(0) and the monetary model can be rejected. Such unit root tests alone enable the authors to reject the monetary model for Denmark, Norway, and Sweden (Table 7, column 3).

Second, Rapach and Wohar (2001a) go on to estimate cointegrating vectors—using 4 different methods—for the following equation:

(24) $$s_t = \beta_0 + \beta_1\left(m_t - m_t^*\right) + \beta_2\left(y_t - y_t^*\right) + \varepsilon_t$$

and test—using 4 different cointegration tests—whether those vectors are consistent with the vector implied by the simple monetary model $\{\beta_1 = 1, \beta_2 = -1\}$(Table 7, columns 4 and 5).[22] All 4 tests find evidence of cointegration for France, Spain, and Italy; 2 tests find evidence for Finland and Portugal; and 1 test finds evidence for Belgium and Switzerland. The estimated coefficients are often close to those implied by the simple monetary model $\{\beta_1 = 1, \beta_2 = -1\}$.

Third, the authors use unit root tests on the residuals from the error correction term $\{s_t - \left(m_t - m_t^*\right) + \left(y_t - y_t^*\right)\}$, which is implied by the simple monetary model. A rejection of the unit root hypothesis is interpreted as supporting the "simple," "long-run" monetary model. The tests produce strong support for the monetary model for the Netherlands, France, Italy, and Spain, moderate support for Belgium, Finland and Portugal and weak support for Switzerland (Table 7, column 6). The authors caution, however, that deviations from monetary fundamentals can be substantial and very persistent.

---

[22] Rapach and Wohar (2001b) defined exchange rates to be the foreign currency price of domestic currency—the inverse of the definition used previously in this paper—and defined the fundamentals as the negative of Mark's fundamentals. For consistency, this paper will use Mark's definitions of the exchange rate and fundamentals. Also, note that Rapach and Wohar (2001b) imposed $\beta_2 = 0$ in their estimate of the cointegrating vector for Finland and Portugal because relative output was found to be I(0) in those cases.

Fourth, a vector error-correction model is estimated to investigate the dynamics of the relation between exchange rates and the fundamentals. The VECM model can be written as follows:

$$(25) \qquad \Delta s_t = \lambda_1 \left(f_{t-1} - s_{t-1} \right) + \sum_{i=1}^{p} a_i \Delta s_{t-i} + \sum_{i=1}^{p} b_i \Delta f_{t-i} + \mu_1 + \varepsilon_{1,t}$$

$$(26) \qquad \Delta f_t = \lambda_2 \left(f_{t-1} - s_{t-1} \right) + \sum_{i=1}^{p} c_i \Delta s_{t-i} + \sum_{i=1}^{p} d_i \Delta f_{t-i} + \mu_2 + \varepsilon_{2,t}$$

Note that fundamentals predict exchange rates in the expected way if either $\lambda_1 > 0$ or $\sum_{i=1}^{P} b_i > 0$.

Similarly, exchange rates predict fundamentals in the expected way if either $\lambda_2 < 0$ or $\sum_{i=1}^{P} c_i > 0$.

Rapach and Wohar (2001a) find that the error correction term $\left(f_{t-1} - s_{t-1} \right)$ predicts exchange rates for Belgium, Finland, and Italy $\left(\lambda_1 > 0, \ \lambda_2 = 0 \right)$. In VECM jargon, monetary fundamentals are said to be *weakly exogenous* for these cases. For Portugal and Spain, the exchange rate is weakly exogenous, the error correction term predicts future fundamental changes but not exchange rate changes $\left(\lambda_1 = 0, \ \lambda_2 < 0 \right)$. The error correction term predicts both future fundamentals and exchange rate changes for France and Switzerland $\left(\lambda_1 > 0, \ \lambda_2 < 0 \right)$. These results are summarized in columns 7 and 8 of Table 7.

Finally, the authors pursue the usual out-of-sample forecasting exercises and find that their Theil U statistics, DM statistics, and *encompassing regressions* show evidence of predictability—beyond the random walk with drift—for Belgium, Italy, and Switzerland, and

some evidence for Finland (from encompassing regressions not shown in Table 7).[23] Not surprisingly, cases in which the VECM showed the exchange rate to be weakly exogenous are out-of-sample forecasting failures. The authors note, however, that forecasting failure can still be consistent with the long-run monetary model if deviations from fundamentals predict future fundamentals (Table 7, column 9).

Figure 1 uses the case in the which the monetary model has the best out-of-sample fit — the case of the U.S. dollar/Italian lira (USD/ITL)—to illustrate how little of the variation in the one-year ahead exchange rate change the monetary model predicts. The top panel shows exchange rate changes and recursive, out-of-sample forecasts from 1939 to the end of the sample in 1995. The bottom panel shows the same data from the beginning of the floating exchange rate era, 1973-1995. The chart appears to show that the monetary model forecasts best in extreme circumstances, such as those seen during the high inflation that Italy suffered during World War II or during 1986-1987, when the dollar weakened again after a period of unusual strength. In almost all periods, however, the monetary model explains very little of the variation in one-year ahead exchange rate changes.

Data revisions

Faust, Rogers, and Wright (2001) examine the impact of yet another issue: data revisions. Previous studies had all assumed that the data they used would be fixed and known to a forecaster. In fact, macroeconomic data such as money supply and output figures are often extensively revised, meaning that data depends on the date on which the series were obtained. In

---

[23] An encompassing regression evaluates the predictive content of forecasting techniques by testing whether realized values of the exchange rate depend positively on predicted values from one or more

other words, if one obtained U.S. output data in April 1992—as the authors infer that Mark did—one might have a different estimate of U.S. output growth for 1991:1 than if one obtained data in April 1993. To investigate the effect of data revisions on forecasting exercises with monetary models, Faust, Rogers, and Wright (2001) obtained 38 data sets, representing the best estimates of the data as it stood on different dates from April 1988 to October 2000.

With these 38 data sets, they first attempted to see if they would obtain the same inference as Mark (1995) by holding back the final 40 quarters of data from each set for an out-of-sample forecasting exercise. That is, each of the 38 data sets were different from each other both because of data revisions and the fact that the final 40 quarters of data would be over different periods. They found that the only vintages of the data that would have produced significant long-horizon predictability were those in a 2-year window around April 1992, the time that Mark collected his data set. The decrease in predictability for other data sets was due to both sample periods—as noted by Kilian (1999)—and data revisions.

To isolate the marginal effect of data revisions, the authors fixed the sample period and compared results using more and less revised data. The more-revised data showed less evidence of predictability. Theil U-statistics rise and p-values fall as the data are revised. For example, data revisions made the 16Q Theil U statistic rise from 0.52 to 0.64 in the German case and from 0.55 to 0.69 in the Japanese case. And the p-values for these cases rise above 0.1. Similarly, the authors estimate a portfolio-balance model using inflation, interest rates and cumulated trade balances—as in Chinn and Meese (1995)—and they find that data revisions have even larger effects than in the monetary model.

forecasting techniques.

Next Faust, Rogers, and Wright (2001) investigate the quality of "real-time" data forecasts. Real-time forecasts use the latest revision of data available at any given point in time to estimate the parameters of the model and make forecasts. That is, real-time exercises can use a different set of data for each forecast period. In contrast, the other forecasting exercises (e.g., Mark, 1995) used one data set—the latest revisions available when the research is done—to estimate the equation and make forecasts. Faust, Rogers, and Wright find that real-time data provides better out-of-sample predictive power—according to out-of-sample relative RMSEs— in almost every case than the latest data revision. Perhaps this should not surprise us. If the exchange rate changes do depend on market participants expectations of future monetary fundamentals (i.e., equation (14)), which are based on currently available (real-time) data, then real-time data should provide better estimates of market expectations.

Finally, the authors find that Federal Reserve forecasts of future variables sometimes outperform actual future values of those independent variables in a MR-type regression multi-period forecast. This is ironic. MR sought to give the monetary models the best possible chance to forecast well by replacing forecasts of future independent variables with actual values. But, (at least some) forecasts of fundamentals predict exchange rates better than the future values of those variables.

The study concludes that both the particular sample period used by Mark (1995) and the particular vintage of data revisions combined to produce better out-of-sample forecasting performance than most data samples and/or data revisions would have. Indeed, only data sets

constructed in about 1992 would have shown evidence of long-horizon predictability. Faust,

Rogers, and Wright (2001) speculate that evidence of forecastability is actually an artifact of data

mining, the tendency to test multiple models on one set of data until, by chance, positive results

are found. Finally, for a given fixed sample, real-time data would have produced better forecasts

than the latest data revisions. Hindsight turns out to handicap the forecasts rather than to

improve them.

Nonlinear models

The monetary model is intuitively appealing but clearly explains very little exchange rate

variability. One explanation for the weak relation is that exchange rates are relatively insensitive

to monetary fundamentals close to equilibrium values but tend to strongly revert to those

fundamentals when the deviation is large. Taylor and Peel (2000), Taylor, Peel, and Sarno

(2001) and Kilian and Taylor (2001) investigate the plausibility of this characterization with

nonlinear models.

Taylor and Peel (2000) estimate a nonlinear error correction model of quarterly exchange

rates and monetary fundamentals for the pound/dollar (GBP/USD) and Deutchesmark/dollar

(DEM/USD) exchange rates from 1973:1 to 1996:4. They find that the exponential smooth

transition autoregressive (ESTAR) model (Granger and Terävirta, 1993) parsimoniously

describes the deviation of the exchange rate from monetary fundamentals. This model predicts

that the exchange rate change will nearly be unpredictable when the deviation from fundamentals

is small, but will strongly revert toward those fundamentals when the deviation is big. The

authors use this to characterize the degree of over- and undervaluation of the exchange rates during the modern period of floating exchange rates. Similarly, Taylor, Peel, and Sarno (2001) show that the same model fits real exchange rates well and explains deviations from purchasing power parity.

Kilian and Taylor (2001) note that a convincing explanation for the nonlinear dynamics of the ESTAR model is lacking. The authors suggest a candidate model in which uncertainty about the fundamental value of the exchange rate deters agents from speculating against small deviations from fundamentals.[24] Monte Carlo studies show that there is more predictability for plausible DGPs at the one- and two-year horizons, so long horizon tests are useful in such an environment. Further, if the ESTAR model is the true DGP, then all past tests of long-horizon predictability are invalid because they assume a linear null data generating process, which is incorrect. Consistent with this prediction, the authors find that in-sample evidence of predictability from seven OECD countries increases "dramatically" with the forecast horizon.[25] Yet, the authors are still unable to find evidence of out-of-sample predictability. They ascribe this to the low power of out-of-sample tests, given the short span of post-Bretton Woods data and the rarity of large departures from fundamentals during that time.

---

[24] Kilian and Taylor (2001) assume that the fundamental value is a function of relative prices rather than money and output, as in the monetary model.

[25] Mark and Sul (2002) find that long-horizon regressions can have asymptotic power advantages over one-period-ahead procedures in cases similar to those found in foreign exchange forecasting. Their Monte Carlo experiments show that the phenomenon might be even more common in finite samples.

WHY DOESN'T THE MONETARY MODEL PREDICT WELL?

One obvious problem is that three of the building blocks of the monetary model, money demand equations, purchasing power parity (PPP), and uncovered interest parity (UIP) do not work very well (Engel, 1996; Engel, 2000). Money demand equations have proven unstable, especially in the United States (Friedman and Kuttner, 1992), but changing the numeraire currency doesn't seem to help the monetary model much.

But that begs the question as to why PPP and UIP perform so poorly. Why are floating exchange rates so volatile and unrelated to prices and interest differentials? Many researchers have claimed that volatile expectations or departures from rationality are likely to account for the failure of exchange rate models. For example, Frankel (1999) argues that exchange rates are detached from fundamentals by swings in expectations about future values of the exchange rate. These fluctuations in exchange rates are essentially bubbles, of the type discussed in Section 2. Four pieces of evidence suggest that expectations are to blame for such behavior: 1) Survey measures of exchange rate expectations are very poor forecasters and the expectations, themselves, are frequently not internally consistent (Frankel and Froot, 1987; Sarno and Taylor 2001); 2) Failure of expectations to be rational is often blamed for the failure of UIP (Engel, 1996); 3) Trend following trading rules appear to make risk-adjusted excess returns, in apparent violation of the efficient markets hypothesis (Neely, 1997; Neely, Weller, and Dittmar, 1997); 4) Switching from a fixed exchange rate to a floating rate—which changes the way expectations are formed—changes the behavior of nominal and real exchange rates and the ability of UIP to explain exchange rate changes.

This latter point requires some explanation. Fixed exchange rates anchor investor sentiment about the future value of a currency because of the government's commitment to stabilize its value. If expectations are based on fundamentals, rather than irrationally changing expectations, then the relationship between fundamentals and exchange rates should be the same under a fixed exchange rate regime as it is under a floating regime. This is not the case. Countries that move from floating exchange rates to fixed exchange rates experience a dramatic change in the relationship between prices and exchange rates. Specifically, real exchange rates (exchange rates adjusted for inflation in both countries) are much more volatile under floating exchange rate regimes, where expectations are not tied down by promises of government intervention (Mussa, 1986). Figure 2 illustrates a typical case: When the German government ceased to fix the DEM to the USD in March 1973, the variability in the real DEM/USD exchange rate increased dramatically. This result suggests that, contrary to the efficient markets hypothesis, swings in investor expectations may detach exchange rates from fundamental values in the short run. Similarly, UIP seems to do such a poor job explaining USD exchange rates while doing a pretty good job with semi-fixed rates such as those found in the EMS (Flood and Rose, 1996). Indeed, Flood and Rose (1999) develop a UIP-based model of the exchange rate that explains why UIP—and exchange rate forecasts—might perform poorly in the short term even with perfect rational agents.

CONCLUSION: THE BIG PICTURE

The seminal work of MR showed that monetary models were unable to forecast exchange rates better than a no-change forecast. Since then, a small army of researchers has attempted to forecast exchange rates with the analytically attractive monetary model. Initial attempts were strikingly unsuccessful; Mark (1995), however, appeared to show that monetary fundamentals could predict exchange rate changes at 3-4 year horizons. Kilian (1999), Berkowitz and Giorgianni (2001), and Faust, Rogers, and Wright (2001) subsequently criticized the underlying assumptions of Mark's study with respect to the stationarity of the data, robustness to sample period, appropriate benchmark for comparison, and the vintage of the data. Attempts to forecast with panel studies (Mark and Sul, 2001; Rapach and Wohar, 2001b) or very long samples (Rapach and Wohar, 2001a) have failed to establish the existence of predictability beyond reasonable doubt.

Other research suggests that exchange rates might be nonlinearly mean reverting to fundamentals and that the intuitively appealing monetary model might therefore provide better predictions when exchange rates deviate substantially from fundamentals (Kilian and Taylor, 2001). Such models also imply that long-horizon regressions might be more informative than short-horizon regressions.

How should one interpret these disparate results? The monetary model is intuitively appealing and monetary variables likely influence exchange rate changes. But, like all models, it simplifies reality. The literature on exchange rate forecasting has shown that the amount of

exchange rate variation explained by monetary models is—at most—small. Further research will doubtless continue to attempt to quantify this predictability. Other future research might profitably explore the way that expectations of asset prices are formed and the factors—like dispersion of belief, risk aversion, and transactions costs—that permit extreme nominal exchange rate variability by hindering arbitrage.

Other conclusions that one might draw depend on the purpose of these forecasting exercises, which is little discussed in the literature. Researchers have usually motivated their work by the desire to evaluate monetary models of the exchange rate (MR) and have cautioned that they are not trying to build the best possible forecasting model (Mark and Sul, 2001), but the negative results for monetary models have nonetheless produced a conventional wisdom in the profession that exchange rate changes cannot be forecast—or cannot be forecast using macroeconomic fundamentals. It is not clear, however, that this is true.

It might seem obvious that both policy makers and firms would want to forecast exchange rates, but it isn't entirely clear why they would wish to do so with monetary fundamentals. Policymakers might wish to forecast exchange rates because of their influence on variables of more direct interest such as output and inflation. But why not directly forecast output and inflation if that is the case? Or, firms might wish to forecast exchange rates to make asset allocation decisions. But that would require a forecast of deviations from UIP, not of exchange rates themselves. Surely the reason for exchange rate forecasts will influence the method of evaluation and the value of those projections. Future research should address this topic.

REFERENCES

Berben, Robert-Paul and van Dijk, Dick. "Does the Absence of Cointegration Explain the
Typical Findings in Long Horizon Regressions?" Unpublished manuscript. Tinbergen
Institute, April 6, 1998.

Berkowitz, Jeremy and Giorgianni, Lorenzo. "Long-Horizon Exchange Rate Predictability?"
Review of Economics and Statistics, February 2001, *83(1)*, pp. 81-91.

Berkowitz, Jeremy and Kilian, Lutz. "Recent Developments in Bootstrapping Time Series."
Econometric Reviews, February 2000, *19(1)*, pp. 1-48.

Bilson, John F.O. "Rational Expectations and the Exchange Rate," in Jacob A. Frenkel and
Henry G. Johnson, eds., *The Economics of Exchange Rates: Selected Studies*, Reading,
MA: Addison-Wesley Press,1978.

Blough, Stephen R. "The Relationship Between Power and Level for Generic Unit Root Tests in
Finite Samples." *Journal of Applied Econometrics*, July-Sept. 1992, *7(3)*, pp. 295-308.

Canova, Fabio. "Modelling and Forecasting Exchange Rates with a Bayesian Time-Varying
Coefficient Model." *Journal of Economic Dynamics and Control*, January-March 1993,
17(1/2), pp. 233-261.

Chinn, Menzie D. and Meese, Richard A. "Banking on Currency Forecasts: How Predictable Is
Change in Money?" *Journal of International Economics*, February 1995, *38(1/2)*, pp.
161-78.

Clarida, Richard H. and Taylor, Mark P."The Term Structure of Forward Exchange Premiums
and The Forecastability of Spot Exchange Rates: Correcting the Errors." *Review of
Economics and Statistics*, 1997, *79(3)*, pp. 353-61.

Clarida, Richard H.; Sarno, Lucio; Taylor, Mark P. and Valente, Giorgio."The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond." *Journal of International Economics,* 2003, (forthcoming).

Diebold, Francis X. and Mariano, Roberto S. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, July 1995, *13(3)*, pp. 253-63.

Dornbusch, Rudiger. "Expectations and Exchange Rate Dynamics." *Journal of Political Economy*, December 1976, *84(6)*, pp. 1161-76.

Engel, Charles. "Long-Run PPP May Not Hold After All." *Journal of International Economics*, August 2000, *51(2)*, pp. 243-73.

Engel, Charles. "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence." *Journal of Empirical Finance*, June 1996, *3(2)*, pp. 123-92.

Evans, M.D.D. and R.K. Lyons. "Order Flow and Exchange Rate Dynamics." Unpublished manuscript, Georgetown University, December 1999.

Fair, Ray C. "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors." *Econometrica*, May 1970, *38(3)*, pp. 507-16.

Fair, Ray C. "Evaluating the Information Content and Money Making Ability of Forecasts from Exchange Rate Equations." Unpublished manuscript, Yale University, May 1999.

Faust, Jon. "Near Observational Equivalence and Unit Root Processes: Formal Concepts and Implications." Board of Governors of the Federal Reserve System International Finance and Economics Discussion Paper: 447, 1993.

Faust, Jon R.; John H. Rogers, and Wright, Jonathan H. "Exchange Rate Forecasting: The Errors We've Really Made." International Finance Discussion Paper 714, December 2001.

Flood, Robert P. and Rose, Andrew K. "Fixes: Of the Forward Discount Puzzle." *Review of Economics and Statistics*, November 1996, *78(4)*, pp.748-52.

Flood, Robert P. and Rose, Andrew K. "Understanding Exchange Rate Volatility Without the Contrivance of Macroeconomics." *Economic Journal*, November 1999, *109(459)*, pp. F660-72.

Frankel, Jeffrey. "How Well Do Foreign Exchange Markets Work: Might a Tobin Tax Help?" in Mahbub ul Haq, Inge Kaul, and Isabelle Grunberg, eds., *The Tobin Tax: Coping with Financial Volatility*. New York and Oxford: Oxford University Press, 1996.

Frankel, Jeffrey A. and Froot, Kenneth A. "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations." *American Economic Review*, March 1987, *77(1)*, pp. 133-53.

Frenkel, Jacob A. "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence." *Scandinavian Journal of Economics*, 1976, *78(2)*, pp. 200-224.

Friedman, Benjamin M. and Kuttner, Kenneth N. "Money, Income, Prices, and Interest Rates." *American Economic Review*, June 1992, *82(3)*, pp. 472-92.

Granger, C.W.J. and Newbold, P. "Spurious Regressions in Econometrics." *Journal of Econometrics*, July 1974, *2(2)*, pp. 111-20.

Granger, Clive W.J. and Teräsvirta, Timo. *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press, 1993.

Groen, Jan J.J. "The Monetary Exchange Rate Model as a Long-Run Phenomenon." *Journal of International Economics*, December 2000, *52(2)*, pp. 299-319.

Groen, Jan J.J. "Long-Run Predictability of Exchange Rates: Is It For Real?" *Empirical Economics*, August 1999, *24(3)*, pp. 451-69.

Hooper, Peter and Morton, John E. "Fluctuations in the Dollar: A Model of Nominal and Real Exchange Rate Determination." *Journal of International Money and Finance*, April 1982, *1*, pp. 39-56.

Horvath, Michael T.K. and Watson, Mark W. "Testing for Cointegration When Some of the Cointegrating Vectors Are Prespecified." *Econometric Theory*, December 1995, *11(5)*, pp. 984-1014.

Johansen, Soren. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica*, November 1991, *59(6)*, pp. 1551-80.

Kendall, M.G. "Note on Bias in the Estimation of Autocorrelation." *Biometrika*, December 1954, 41(3/4), pp. 403-4.

Kilian, Lutz. "Exchange Rates and Monetary Fundamentals: What Do We Learn From Long-Horizon Regressions?" *Journal of Econometrics*, September/October 1999, *14(5)*, pp. 491-510.

Kilian, Lutz and Taylor, Mark. "Why Is It So Difficult to Beat Random Walk Forecasts of Exchange Rates?" Unpublished manuscript, October 2001.

Lane, Philip R. "The New Open Economy Macroeconomics: A Survey." *Journal of International Economics*, August 2001, *54(2)*, pp. 235-66.

Levin, Andrew and Lin, Chien-Fu. "Unit Root Tests in Panel Data: New Results." University of California, San Diego Department of Economics Working Paper: 93-56, December 1993.

Lucas, Robert E. "Interest Rates and Currency Prices in a Two-Country World." *Journal of Monetary Economics*, November 1982, *10(3)*, pp. 335-59.

Mankiw, N. Gregory and Shapiro, Matthew D. "Do We Reject Too Often? Small Sample

  Properties of Tests of Rational Expectations Models." *Economics Letters,* 1986, *20*, pp.

  139-45.

Mark, Nelson C. *International Macroeconomics and Finance: Theory and Econometric*

  Methods. Malden, MA: Blackwell Publishers, 2001.

Mark, Nelson C. "Exchange Rates and Fundamentals: Evidence on Long-Horizon

  Predictability." *American Economic Review*, March 1995, *85(1)*, pp. 201-18.

Mark, Nelson C. and Sul, Donggyu. "Nominal Exchange Rates and Monetary Fundamentals:

  Evidence from a Small Post-Bretton Woods Panel." *Journal of International Economics*,

  February 2001, *53(1)*, pp. 29-52.

_____ and _____. "Asymptotic Power Advantages of Long Horizon

  Regressions." Unpublished manuscript, Ohio State University, April 2002.

Mark, Nelson C.; Ogaki, Masao and Sul, Donggyu. "Dynamic Seemingly Unrelated

  Cointegrating Regression." Unpublished manuscript, Ohio State University, July 5, 2000.

Marriott, F.H.C. and Pope, J.A. "Bias in the Estimation of Autocorrelations." *Biometrika*,

  December 1954, *41(3/4)*, pp. 390-402.

Meese, Richard A. "Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates?"

  Journal of Political Economy, April 1986, *94(2)*, pp. 345-73.

Meese, Richard A. and Rogoff, Kenneth. "Empirical Exchange Rate Models of the Seventies: Do

  They Fit Out of Sample?" *Journal of International Economics*, February 1983a, *14(1/2)*,

  pp. 3-24.

_____ and _____. "The Out-of-Sample Failure of Empirical Exchange Rate

Models: Sampling Error or Misspecification?" in Jacob Frenkel, ed., *Exchange Rates and*

International Macroeconomics. Chicago: NBER and University of Chicago Press, 1983b.

Meese, Richard A. and Rose, Andrew K. "Nonlinear, Nonparametric, Nonessential Exchange

Rate Estimation." *American Economic Review*, May 1990, *80(2)*, pp.192-96.

Mussa, Michael. "The Exchange Rate, the Balance of Payments and Monetary and Fiscal Policy

under a Regime of Controlled Floating." *Scandinavian Journal of Economics*, 1976,

78(2), pp. 229-48.

Mussa, Michael. "Empirical Regularities in the Behavior of Exchange Rates and Theories of the

Foreign Exchange Market," in Karl Brunner and Allan H. Meltzer, eds., *Carnegie-*

Rochester Conference Series on Public Policy: Policies for Employment, Prices and

Exchange Rates, Vol. 11, 1979.

Mussa, Michael. "Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates:

Evidence and Implications," in Karl Brunner and Allan H. Meltzer, eds., *Carnegie-*

Rochester, Conference Series on Public Policy: Real Business Cycles, Real Exchange

Rates, and Actual Policies, Vol. 25, Autumn 1986, pp. 117-214.

Neely, Christopher J. "Technical Analysis in the Foreign Exchange Market: A Layman's Guide."

Federal Reserve Bank of St. Louis *Review*, September/October 1997, *79(5)*, pp. 23-38.

Neely, Christopher; Weller, Paul; and Dittmar, Rob. "Is Technical Analysis in the Foreign

Exchange Market Profitable? A Genetic Programming Approach." *Journal of Financial*

and Quantitative Analysis, December 1997, *32(4)*, pp. 405-26.

Pesaran, M. Hashem; Shin, Yongcheol and Smith, Ron P. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association*, June 1999, *94(446)*, pp. 621-34.

Phillips, Peter C. B. "Understanding Spurious Regressions in Econometrics." *Journal of Econometrics*, December 1986, *33(3)*, pp. 311-40.

Rapach, David E. and Wohar, Mark E. "Testing the Monetary Model of Exchange Rate Determination: New Evidence from a Century of Data." Unpublished manuscript, November 2001a, *Journal of International Economics* (forthcoming).

Rapach, David E. and Wohar, Mark E. "Testing the Monetary Model of Exchange Rate Determination: A Closer Look at Panels." September 2001b.

Rudebusch, Glenn D. "The Uncertain Unit Root in Real GNP." *American Economic Review*, March 1993, *83(1)*, pp. 264-72.

Sarno, Lucio. "Toward a New Paradigm in Open Economy Modeling: Where Do We Stand?" Federal Reserve Bank of St. Louis *Review*, May/June 2001, *83(3)*, pp. 21-36.

Sarno, Lucio and Taylor, Mark P. "The Microstructure of the Foreign Exchange Market: A Selective Survey of the Literature." *Princeton Studies in International Economics*, No. 89, April 2001.

Sarno, Lucio and Taylor, Mark P. *The Economics of Exchange Rates*. Cambridge and New York: Cambridge University Press, 2002 (forthcoming).

Schinasi, Garry J. and Swamy, P.A.V.B. "The Out-of-Sample Forecasting Performance of Exchange Rate Models When Coefficients Are Allowed to Change." *Journal of International Money and Finance*, September 1989, *8(3)*, pp. 375-90.

Stambaugh, Robert F. "Bias In Regressions With Lagged Stochastic Regressors." Working Paper

  No. 156, Center for Research in Security Prices, January 1986.

Stock, James H. and Watson, Mark W. "A Simple Estimator of Cointegrating Vectors in Higher

  Order Integrated Systems." *Econometrica*, July 1993, 61(4), pp. 783-820.

Stockman, Alan C. "A Theory of Exchange Rate Determination." *Journal of Political Economy*,

  August 1980, *88(4)*, pp. 673-98.

Taylor, Mark P. and Peel, David A. "Nonlinear Adjustment, Long-Run Equilibrium and

  Exchange Rate Fundamentals." *Journal of International Money and Finance*, February

  2000, *19(1)*, pp. 33-53.

Taylor, Mark P.; Peel, David A. and Sarno, Lucio. "Nonlinear Mean-Reversion in Real

  Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles."

  International Economic Review 2001, *42(4),* pp. 1015-42.

Wolff, Christian C.P. "Time-Varying Parameters and the Out-of-Sample Forecasting

  Performance of Structural Exchange Rate Models." *Journal of Business and Economic*

  Statistics, January 1987, *5(1)*, pp. 87-97.

Wolff, Christian C.P. "Exchange Rates, Innovations and Forecasting." *Journal of International*

  Money and Finance, March 1988, *7(1)*, pp. 49-61.

Figure 1: Forecasts of the Change in the ITL/USD Exchange Rate
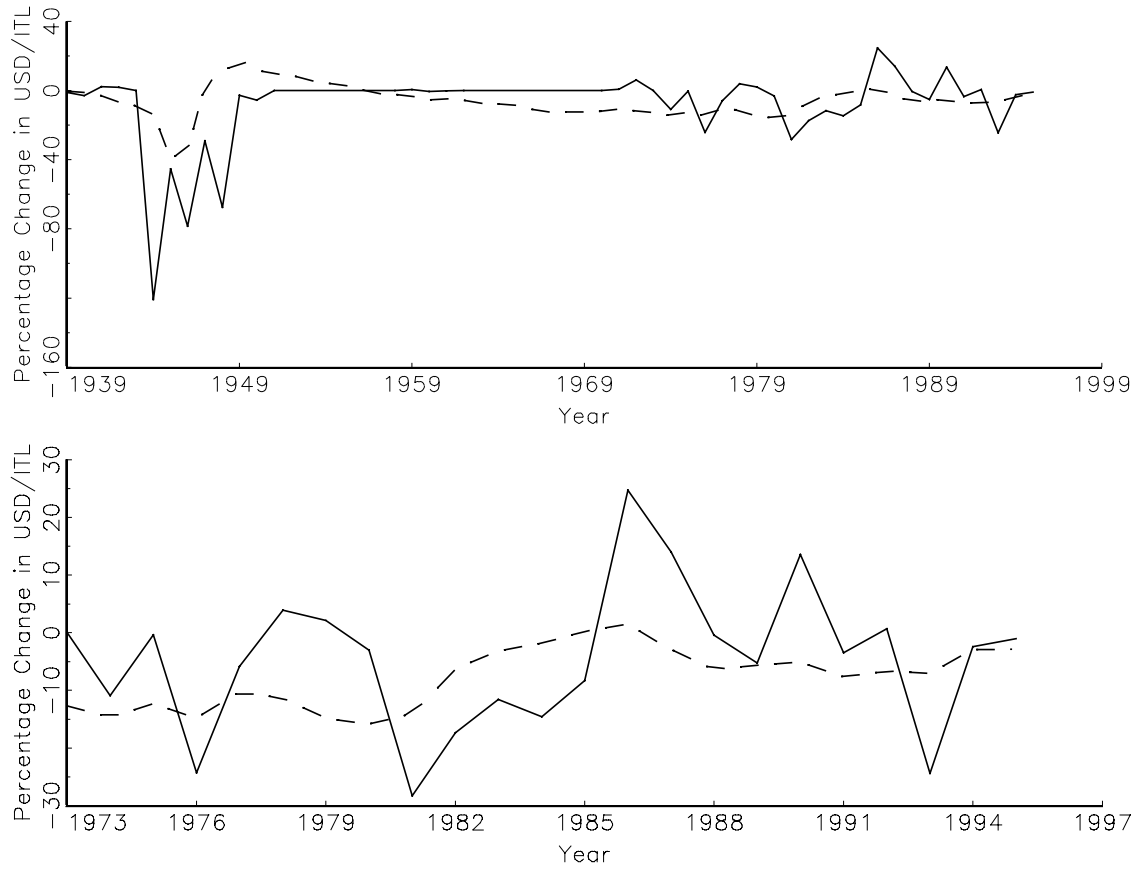
NOTE: The upper panel depicts the percentage annual change in the USD/ITL exchange rate (solid line) and the predicted value from the monetary model (dashed line) from 1939 to 1995. The lower panel depicts the same data over the floating rate period, 1973-1995.

Figure 2: Monthly Percentage Changes in the DEM/USD Real Exchange Rate.
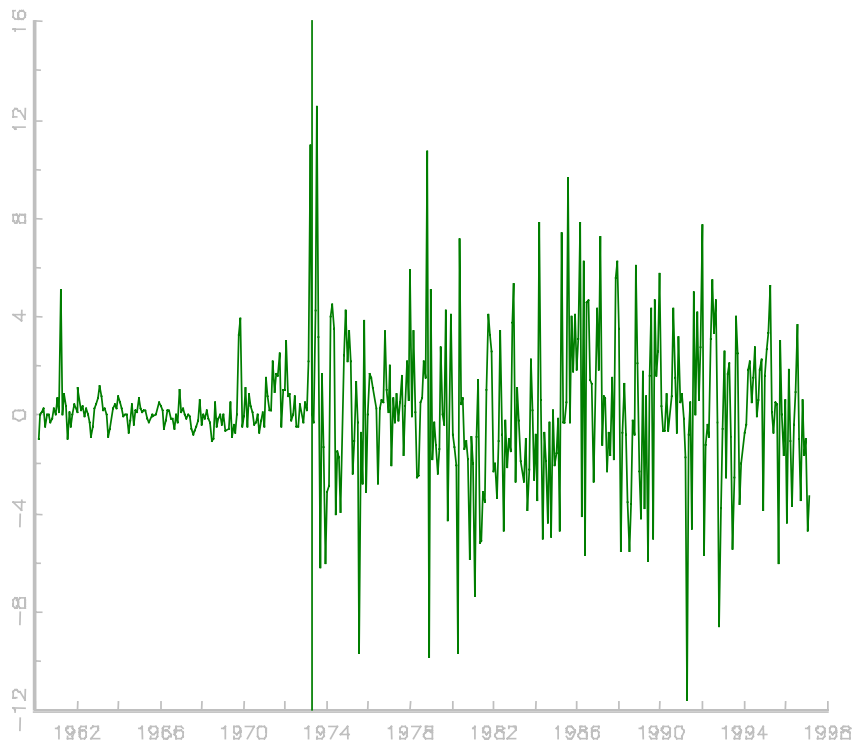
NOTE: Real exchange rates became much more volatile after the end of the Bretton-Woods System of fixed exchange rates in March 1973. The vertical line denotes this break date in the series.

Boxed Insert: A Short Primer on Persistence and Cointegration

A data series is said to be *stationary* if neither its mean nor any of its autocovariances depend on the date t. An implication of this is that the series will tend to be mean-reverting, to tend to return to its expected value when it departs from it. The longer such a series takes to return to its mean, the more *persistent* it is said to be. A series that is infinitely persistent will never tend to return to any mean value—its expected value doesn't exist; such a series is *nonstationary*.[26] Note that if a variable X_t has zero mean and is uncorrelated at all leads and lags—i.e., it is *white noise*—then X_t is stationary. If we define Y_t as the cumulated sums of X_t

$(Y_t = \sum_{i=0}^{t} X_i$), then, as $t \to \infty$, the variance of Y_t becomes infinite, Y_t does not tend to revert to any

value, so it is nonstationary. Y_t is also said to be I(1), *integrated of order one*—because it is the partial sum (or integral) of the X_is—and X is referred to as I(0). Y_t is also sometimes called a *random walk* because changes to it are unpredictable (random) from other

information, $Y_t = Y_{t-1} + X_t$, $E(\Delta Y_t) = 0$. Note that if X_t had a constant non-zero mean, that is

$X_t = \delta + \varepsilon_t$, where ε_t is white noise, then Y_t would be *random walk with drift:* $Y_t = Y_{t-1} + \delta + \varepsilon_t$,

$E(\Delta Y_t) = \delta$.

If at least one linear combination of I(1) variables is I(0), the variables are said to be *cointegrated*. Economic theory often implies that a linear combination of variables will be stationary. For example, one version of purchasing power parity says that exchange rates should be cointegrated with relative price levels—assuming that both are I(1) variables.

The persistence and cointegration of variables can have important implications for specifying regressions. For example, a regression equation only makes sense if the dependent and independent variables can be written so that the error is I(0). This requires that either both sides are I(0) or some combination of them is I(0). If the error term is I(1), its mean does not exist, coefficient estimates will be inconsistent and forecasts will be biased. Generally, regressions are valid only if coefficients exist that make the error term stationary (Phillips,1986).

A bivariate vector autoregression with a cointegrating relation between I(1) variables can be written in vector-error correction (VECM) form as follows:

$$\Delta s_t = \lambda_1 \left(f_{t-1} - s_{t-1} \right) + \sum_{i=1}^{p} a_i \Delta s_{t-i} + \sum_{i=1}^{p} b_i \Delta f_{t-i} + \mu_1 + \varepsilon_{1,t}$$

$$\Delta f_t = \lambda_2 \left(f_{t-1} - s_{t-1} \right) + \sum_{i=1}^{p} c_i \Delta s_{t-i} + \sum_{i=1}^{p} d_i \Delta f_{t-i} + \mu_2 + \varepsilon_{2,t}$$

Under the null hypothesis of no cointegration between the variables, the individual variables are I(1), and a VAR in differences is appropriate. Such a VAR implies that $\lambda_1 = 0$ and $\lambda_2 = 0$. To test this hypothesis, Horvath and Watson (1995) suggest a Wald test of the null that $\lambda_1 = 0$ and $\lambda_2 = 0$ in the above representation.

Unfortunately, it is difficult to tell the difference between variables that are I(1)—having no mean—and those that are I(0) but highly persistent, taking a long time to return to their mean. Therefore, resolving the question of whether fundamentals are cointegrated with exchange rates

---

[26] If the expected value of a series does not exist, it is nonstationary. But nonstationary series can have expected values.

is more difficult than one might think. In fact, it can be shown that stationary and nonstationary processes are *observationally equivalent* (Blough, 1992; Faust,1993). In other words, there is a nonstationary process that will behave arbitrarily close to any given stationary process, and vice versa. In practice, it is frequently the case that a series of interest could be consistent with either stationarity or nonstationarity.

The problem of *spurious regression* occurs when an I(1) variable is regressed on an unrelated I(1) variable. Spurious regressions, investigated by Granger and Newbold (1974) and analyzed further by Phillips (1986), are regressions that relate independent random walks and produce apparently statistically significant—but inconsistent and meaningless—coefficients. Intuitively, spurious regressions occur because the moment matrix of the regressors (X'X) doesn't converge to anything and so becomes arbitrarily large, making the reported standard error of the coefficient too small and permitting false rejections of the hypothesis that the coefficient is equal to zero.

Even a variable that is not I(1) can cause problems in a regression if it is *persistent*—highly autocorrelated. It has long been understood that if the independent variable in a regression is persistent, the coefficient estimates in that regression will be biased. Marriott and Pope (1954) and Kendall (1954) studied the phenomenon in the autoregressive case and Mankiw and Shapiro (1986) and Stambaugh (1986) extended the results to multiple equation models. To understand why persistence creates finite sample bias, consider the case of an AR(1) process:

$y_t = \beta y_{t-1} + \varepsilon_t$ where $y_0 = 0$ and $E(\varepsilon_t^2) = \sigma^2$ and $E(\varepsilon_t) = 0$. The OLS estimator is given by

$$\hat{\beta} = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1} y_{t-1}} = \frac{\sum_{t=1}^{T} y_{t-1} (\beta y_{t-1} + \varepsilon_t)}{\sum_{t=1}^{T} y_{t-1} y_{t-1}}$$ and the difference between $\hat{\beta}$ and β is given by:

$$\hat{\beta} - \beta = \frac{\sum_{t=1}^{T} y_{t-1} \varepsilon_t}{\sum_{t=1}^{T} y_{t-1}^2}.$$ With persistent regressors, the expectation of $\hat{\beta} - \beta$ is not equal to zero

because there is positive correlation between the product $y_{t-1}\varepsilon_t$ and y_{t+j}^2 for $j \geq 0$. In other words,

if the product of the regressor (y_{t-1}) and the error term (ε_t) is large and positive, then they are

probably both large and of the same sign. In this case, near-term future values of the series (y_t)

will tend to be far from the unconditional mean and their squares will be large too. Conversely,

if $y_{t-1}\varepsilon_t$ is negative, then y_t will tend to be closer to its mean (0) than y_{t-1} and so y_t^2 will tend to be

small. Thus an estimator in a regression with a persistent regressor will be biased in finite

samples.

Table 1: Mark's (1995) results using the DGP described in equations (15) and (16), 1973:2-1991:4.

Country	Horizon	Beta	Adj-Beta	t(20)	p-value	R^2	p-value	OUT/RW	p-value	DM(20)	p-value
Canada	1	0.040	0.029	3.051	**0.070**	0.059	**0.058**	0.998	0.181	0.061	0.184
	4	0.155	0.109	2.389	0.183	0.179	**0.090**	1.119	0.537	-1.270	0.472
	8	0.349	0.258	2.539	0.215	0.351	**0.065**	1.145	0.388	-1.036	0.361
	12	0.438	0.317	1.961	0.340	0.336	0.150	1.436	0.550	-1.916	0.531
	16	0.450	0.286	1.542	0.443	0.254	0.305	1.699	0.615	-2.596	0.542
Germany	1	0.035	0.011	1.836	0.291	0.015	0.419	1.015	0.340	-0.932	0.403
	4	0.205	0.106	2.902	0.181	0.104	0.267	1.037	0.289	-1.345	0.506
	8	0.554	0.363	3.487	0.191	0.265	0.178	1.002	0.226	-0.027	0.225
	12	0.966	0.676	6.329	**0.069**	0.527	**0.060**	0.796	0.109	4.246	**0.058**
	16	1.324	0.955	9.256	**0.038**	0.762	**0.015**	0.524	**0.036**	8.719	**0.045**
Japan	1	0.047	0.012	1.396	0.398	0.020	0.332	0.988	0.248	1.571	0.137
	4	0.263	0.132	2.254	0.278	0.125	0.205	0.928	0.210	2.302	0.121
	8	0.575	0.315	3.516	0.209	0.301	0.126	0.819	0.170	3.096	0.109
	12	0.945	0.564	4.889	0.152	0.532	**0.036**	0.712	0.149	3.319	0.146
	16	1.273	0.790	4.919	0.169	0.694	**0.011**	0.574	0.121	5.126	0.157
Switzerland	1	0.074	0.044	2.681	0.125	0.051	**0.096**	0.997	0.266	0.066	0.282
	4	0.285	0.167	3.248	0.148	0.180	**0.091**	0.981	0.256	0.218	0.265
	8	0.568	0.336	4.770	**0.095**	0.336	**0.077**	0.917	0.219	0.703	0.240
	12	0.837	0.509	8.013	**0.024**	0.538	**0.026**	0.738	0.122	2.933	0.135
	16	1.086	0.672	17.406	**0.001**	0.771	**0.001**	0.411	**0.026**	9.650	**0.071**

NOTE: The table was constructed using programs and data supplied by Nelson Mark. The null DGP constructs the exchange rate as a random walk with drift, the error correction term, (f_t-s_t), is constructed to follow an AR(p) process, and errors to equations (15) and (16) are drawn with nonparametric bootstrapping. The benchmark for out-of-sample forecast comparison is the driftless random walk, a no-change forecast. The column denoted "Beta" denotes the estimate of β_k from equation (14); "Adj-Beta" denotes the estimate of β_k adjusted for endogenous regressor bias; "t(20)" is the t-statistic computed using a 20-period window for Newey-West corrected standard errors; "R^2" is the R^2 of the forecast equation (14); "OUT/RW" is the ratio of the forecast RMSE to the RMSE of the no-change benchmark; "DM(20)" is the Diebold-Mariano statistic, computed using a 20-period window for construction of the covariance matrix for the test of equality between the RMSEs of the regression forecast and the no-change benchmark forecast. Boldface p-values denote results significant at the 10 percent level.

Table 2: Results from a Horvath-Watson test for cointegration (1973:2-1994:4).

	λ_1	t-statistic	λ_2	t-statistic	Wald Statistic	Lag Length
Canada	0.031	1.60	-0.016	-1.48	5.801	2
Germany	0.033	1.11	0.001	0.11	1.233	2
Japan	0.057	1.49	-0.002	-0.25	2.442	3
Switzerland	0.079	1.98	-0.014	-1.78	7.842	2

NOTE: This table reproduces Table 7 from Berkowitz and Giorgianni (2001). It shows the results of a Horvath-Watson test for cointegration. Wald statistics exceeding the 10 percent critical value of 6.63 or the 5 percent critical value of 8.47 reject the null of no cointegration.

Table 3: Results of a forecasting exercise using the DGP described in equations (17) and (18), 1973:2-1991:4

Country	Horizon	Beta	t(20)	p-value	R^2	OUT/RW	p-value	DM(20)	p-value
Canada	1	0.040	3.051	**0.095**	0.059	0.998	0.405	0.061	**0.441**
	4	0.155	2.398	0.217	0.179	1.119	0.412	-1.270	0.849
	8	0.349	2.539	0.225	0.351	1.145	0.712	-1.036	0.958
	12	0.438	1.961	0.352	0.336	1.436	0.317	-1.916	0.592
	16	0.450	1.542	0.458	0.254	1.699	0.196	-2.596	0.466
Germany	1	0.035	1.836	0.510	0.015	1.015	0.969	-0.932	0.724
	4	0.205	2.902	0.354	0.104	1.037	0.914	-1.345	0.522
	8	0.554	3.487	0.354	0.265	1.002	0.809	-0.027	0.814
	12	0.966	6.329	0.165	0.527	0.796	0.406	4.246	**0.093**
	16	1.324	9.256	**0.096**	0.762	0.524	0.113	8.719	**0.030**
Japan	1	0.047	1.396	0.516	0.020	0.988	0.477	1.571	0.286
	4	0.263	2.254	0.353	0.125	0.928	0.396	2.302	0.215
	8	0.575	3.516	0.228	0.301	0.819	0.304	3.096	0.172
	12	0.945	4.889	0.166	0.532	0.712	0.233	3.319	0.173
	16	1.273	4.919	0.216	0.694	0.574	0.142	5.126	0.109
Switzerland	1	0.074	2.681	0.210	0.051	0.997	0.642	0.066	0.704
	4	0.285	3.248	0.181	0.180	0.981	0.596	0.218	0.676
	8	0.568	4.770	**0.095**	0.336	0.917	0.458	0.703	0.621
	12	0.837	8.013	**0.032**	0.538	0.738	0.214	2.933	0.203
	16	1.086	17.410	**0.006**	0.771	0.411	**0.026**	9.650	*0.019*

NOTE: This is excerpted from Table 4 in BG. Errors for the null DGP described in equations (17) and (18) were drawn with a parametric bootstrap. The out-of-sample benchmark for comparison is a random walk with drift. See the notes to Table 1 for column headings. Boldface p-values denote results significant at the 10 percent level. Shaded boxes indicate results that would be significant with the cointegration assumption but are no longer significant without it. Italicized p-values indicate figures that were not significant with the cointegration assumption but are now significant without it.

Table 4: Results of a forecasting exercise using the DGP described in equations (17) and (18), an extended sample, 1973:2-1994:4

Country	Horizon	Beta	t(20)	p-value	R²	OUT/RW	p-value	DM(20)	p-value
Canada	1	0.035	3.013	0.136	0.041	0.994	0.428	0.169	0.540
	4	0.147	2.475	0.295	0.155	1.040	0.853	-0.270	0.724
	8	0.336	2.489	0.325	0.331	1.078	0.818	-0.316	0.756
	12	0.430	1.799	0.505	0.334	1.280	0.416	-0.842	0.967
	16	0.441	1.350	0.606	0.236	1.542	0.260	-1.580	0.782
Germany	1	0.038	2.269	0.431	0.021	0.998	0.715	0.117	0.713
	4	0.156	2.369	0.487	0.082	1.005	0.821	-0.124	0.836
	8	0.396	2.617	0.523	0.216	1.055	0.987	-0.286	0.930
	12	0.697	3.250	0.474	0.393	1.133	0.886	-0.340	0.961
	16	1.019	3.956	0.432	0.601	1.235	0.720	-0.518	0.967
Japan	1	0.032	1.079	0.745	0.012	0.976	0.242	1.818	0.267
	4	0.174	1.360	0.723	0.065	0.942	0.429	0.991	0.490
	8	0.422	2.093	0.609	0.182	0.895	0.485	0.986	0.587
	12	0.719	3.027	0.485	0.364	0.932	0.715	0.338	0.820
	16	0.907	3.050	0.533	0.486	1.067	0.855	-0.235	0.990
Switzerland	1	0.080	2.559	0.304	0.052	0.949	*0.035*	2.195	0.180
	4	0.287	3.195	0.250	0.175	0.838	*0.063*	1.629	0.305
	8	0.566	4.635	0.165	0.332	0.722	*0.076*	1.222	0.475
	12	0.848	7.942	**0.066**	0.538	0.455	*0.015*	1.428	0.505
	16	1.081	18.820	**0.007**	0.769	0.347	**0.007**	1.614	0.568

NOTE: This is excerpted from Table 5 in BG. Errors for the null DGP described in equations (17) and (18) were drawn with a parametric bootstrap. The out-of-sample benchmark for comparison is a random walk with drift. Boldface p-values denote results significant at the 10 percent level. Shaded boxes indicate results that were significant for the original sample (1973:2-1991:4), but are no longer significant with the extended sample. Italicized p-values indicate the reverse, results that were not significant for the original sample (1973:2-1991:4), but now are significant with the extended sample.

Table 5: Results of a forecasting exercise using the BG DGP #2, an unrestricted VAR, equations (19) and (20), 1973:2-1994:4

Country	Horizon	Beta	t(20)	p-value	R²	OUT/RW	p-value	DM(20)	p-value
Canada	1	0.035	3.013	0.426	0.041	0.994	0.557	0.169	0.597
	4	0.147	2.475	0.873	0.155	1.040	0.832	-0.270	0.799
	8	0.336	2.489	0.957	0.331	1.078	0.845	-0.316	0.827
	12	0.431	1.799	0.989	0.334	1.284	0.907	-0.842	0.877
	16	0.440	1.352	0.988	0.236	1.542	0.946	-1.580	0.914
Germany	1	0.038	2.269	0.740	0.021	0.998	0.497	0.117	0.483
	4	0.156	2.369	0.765	0.082	1.005	0.534	-0.124	0.547
	8	0.396	2.617	0.759	0.216	1.055	0.590	-0.286	0.570
	12	0.697	3.255	0.691	0.393	1.133	0.661	-0.340	0.585
	16	1.019	3.956	0.614	0.601	1.235	0.730	-0.518	0.636
Japan	1	0.032	1.079	0.986	0.012	0.976	0.359	1.818	0.214
	4	0.174	1.365	0.982	0.065	0.942	0.504	0.991	0.433
	8	0.422	2.093	0.938	0.182	0.895	0.512	0.986	0.475
	12	0.719	3.027	0.855	0.364	0.932	0.583	0.338	0.604
	16	0.907	3.050	0.843	0.486	1.067	0.714	-0.235	0.692
Switzerland	1	0.080	2.559	0.513	0.052	0.949	**0.062**	2.195	0.150
	4	0.287	3.195	0.457	0.175	0.838	0.103	1.629	0.260
	8	0.566	4.635	0.300	0.332	0.722	**0.085**	1.222	0.330
	12	0.848	7.942	**0.090**	0.538	0.455	**0.010**	1.428	0.301
	16	1.081	18.820	**0.001**	0.769	0.347	**0.004**	1.614	0.300

NOTE: This is excerpted from Table 6 in BG. Errors for the null DGP described in equations (19) and (20) were drawn with a parametric bootstrap. The out-of-sample benchmark for comparison is a random walk with drift. See the notes to Table 1 for column headings. Boldface p-values denote results significant at the 10 percent level. Shaded cells indicate results that were significant in Table 4, but are no longer significant with the unrestricted VAR as the DGP.

Christopher J. Neely and Lucio Sarno

Table 6: Results of a forecasting exercise using Kilian's DGP, equations (21) and (22), 1973:2-1994:4.

| Country | Horizon | benchmark: driftless RW | | | benchmark: RW with drift | | |
		t(20) p-value	OUT/RW p-value	DM(20) p-value	t(20) p-value	OUT/RW p-value	DM(20) p-value
Canada	1	**0.054**	**0.045**	**0.065**	**0.054**	**0.048**	**0.055**
	4	0.132	0.147	0.139	0.131	**0.055**	**0.050**
	8	0.172	0.139	0.145	0.172	**0.052**	**0.022**
	12	0.286	0.341	0.247	0.287	0.129	0.116
	16	0.392	0.539	0.469	0.391	0.250	0.345
Germany	1	0.420	0.348	0.470	0.346	0.352	0.537
	4	0.486	0.329	0.675	0.395	0.309	0.407
	8	0.520	0.334	0.316	0.402	0.308	0.303
	12	0.473	0.268	0.249	0.359	0.272	0.274
	16	0.428	0.206	0.231	0.344	0.281	0.299
Japan	1	0.771	**0.080**	**0.042**	0.629	0.269	0.360
	4	0.738	**0.079**	**0.071**	0.589	0.276	0.290
	8	0.620	**0.072**	**0.083**	0.473	0.217	0.221
	12	0.497	**0.099**	0.142	0.362	0.276	0.277
	16	0.557	0.112	0.173	0.403	0.389	0.346
Switzerland	1	0.150	0.107	0.123	0.150	**0.061**	**0.081**
	4	0.132	0.117	0.130	0.132	**0.091**	0.102
	8	**0.095**	0.113	0.137	**0.095**	**0.089**	0.112
	12	**0.036**	**0.046**	**0.092**	**0.035**	**0.032**	**0.064**
	16	**0.001**	**0.003**	**0.073**	**0.001**	**0.007**	**0.049**

NOTE: This table is excerpted from Tables 2 and 3 in the December 1997 working paper version of Kilian (1999). The left-hand panel uses a driftless random walk for the DGP and out-of-sample benchmark. The right-hand panel uses a random walk with drift for the DGP and out-of-sample benchmark. The columns display the p-values for the t-statistic with 20 period window, ratio of RMSEs, and Diebold-Mariano statistics for differences in RMSEs, respectively. Boldface p-values denote results significant at the 10 percent level.

Christopher J. Neely and Lucio Sarno

Table 7: Summary of Rapach and Wohar (2001a)

Country	Sample	Unit Root Tests on $\{s_t, m_t-m_t^*, y_t-y_t^*\}$	DOLS cointegrating vector	Tests that Support Cointegration	Unit Root Tests on $s_t -[m_t-m_t^*]+[y_t-y_t^*]$	$(f_t - s_t)$ predicts Δs_{t+1} in a VECM	$(f_t - s_t)$ predicts Δf_{t+1} in a VECM	4Q Ahead Theil U statistics
Australia	1880-1995		{0.45, -0.19}					
Belgium	1880-1989		{1.00, -1.01}	1 of 4	Moderate support	Yes	No	0.98
Canada	1880-1995		{0.13, 0.10}					
Denmark	1885-1995	Reject the model						
Finland	1911-1995		{1.01, NA}	3 of 4	Moderate support	Yes	No	1.02
France (trend)	1880-1989		{1.03, -1.16}	4 of 4	Strong support	Yes	Yes	1.02
Italy	1880-1995		{0.96, -1.34}	4 of 4	Strong support	Yes	No	0.94
Netherlands	1900-1992		NA	NA	Strong support			
Norway	1899-1995	Reject the model						
Portugal	1929-1995		{1.07, NA}	3 of 4	Moderate support	No	Yes	1.01
Spain	1901-1995		{0.86, -1.29}	4 of 4	Strong support	No	Yes	1.03
Sweden	1880-1995	Reject the model						
Switzerland (trend)	1880-1995		{0.86, -1.30}	1 of 4	Weak support	Yes	Yes	0.99
United Kingdom	1880-1995		{0.45, -0.99}					

Notes: This table summarizes the results from Rapach and Wohar (2001a). "Trend" after a country name indicates that a trend was permitted in cointegrating relations. Column 2 displays the sample period. Column 3 displays "Reject" if univariate unit root tests on the 3 series $\{s_t, m_t-m_t^*, y_t-y_t^*\}$ permit one to reject that a linear combination of them can be stationary. Rejecting this hypothesis permits one to immediately reject the monetary model of exchange rates. Column 4 shows dynamic OLS estimates of the cointegrating vector (Stock and Watson, 1993). The monetary model implies a vector of $\{1, -1\}$. NA in a cell indicates that either all variables were stationary (the Netherlands) or that relative output $(y-y^*)$ is stationary (Finland and Portugal). Column 5 provides the number of cointegration tests (out of 4) that support the hypothesis that exchange rates are cointegrated with fundamentals. Column 6 shows the degree to which tests on $s_t -[m_t-m_t^*]+[y_t-y_t^*]$ reject the null of a unit root and thereby support the monetary model. Columns 7 and 8 show whether the coefficients λ_1 and λ_2 from the VECM described by equations (25) and (26) are statistically significant, respectively. Column 9 provides the one-year ahead out-of-sample Theil U statistic in each case. Statistics less than one indicate that the monetary model outforecast the random walk benchmark.

Christopher J. Neely and Lucio Sarno