Can Markov Switching Models Predict Excess Foreign Exchange Returns?

Michael J. Dueker
and
Christopher J. Neely

_____

Can Markov switching models predict excess foreign exchange returns?

Michael Dueker[†]

Christopher J. Neely[*]

January 30, 2006

Abstract: This paper merges the literature on technical trading rules with the literature on Markov switching to develop economically useful trading rules. The Markov models' out-of-sample, excess returns modestly exceed those of standard technical rules and are profitable over the most recent subsample. A portfolio of Markov and standard technical rules outperforms either set individually, on a risk-adjusted basis. The Markov rules' high excess returns contrast with mixed performance on statistical tests of forecast accuracy. There is no clear source for the trends, but permitting the mean to depend on higher moments of the exchange rate distribution modestly increases returns.

Keywords: technical trading rules, Markov switching, exchange rates, excess returns, predictability

JEL subject numbers: F31, G15

† Assistant Vice-President, Research Department
  Federal Reserve Bank of St. Louis
  P.O. Box 442 St. Louis, MO 63166
  (314) 444-8594, (314) 444-8731 (fax)
  mdueker@stls.frb.org

* Research Officer, Research Department
  Federal Reserve Bank of St. Louis
  P.O. Box 442 St. Louis, MO 63166
  (314) 444-8568, (314) 444-8731 (fax)
  neely@stls.frb.org

---

1.   Introduction

Research on exchange rate predictability has revealed a disconcerting pair of facts: 1)

Exchange rate changes cannot be forecast with fundamentals at horizons of less than a year

(Meese and Rogoff, 1983), and 2) Trend-following technical analysis—the use of past prices to

make trading decisions—would have been profitable on dollar exchange rates over extended

periods of time (Sweeney, 1986; Levich and Thomas,1993; Neely, Weller and Dittmar, 1997;

Chang and Osler, 1999; LeBaron, 1999; Okunev and White, 2003; Olson, 2004). Almost all of

the technical rules studied have been variants of moving-average or filter rules, which provide

buy-sell decisions, rather than rules derived from time series models, even though the latter could

potentially provide much more information in the form of a full density forecast.

Given that trend-following trading rules have shown excellent predictive content measured

by the excess return standard, one naturally considers whether ARIMA or Markov models, the

workhorses of univariate time series analysis, could imply this same predictive content.

Dewachter (2001) argues that ARIMA models may have been suboptimal generators of trading

signals. Estimating an ARIMA model on data generated by a typical Markov model leads to the

estimation of an ARIMA (1,0,2) in which the first AR and first MA coefficients nearly cancel,

resulting in some loss of information. Markov switching models explicitly estimate time-varying

moments, naturally modeling exchange rate returns in a changing risk environment. Such a

model exploits the advantages of statistical techniques to construct trading rules.

The use of Markov switching models to forecast and trade exchange rates at the daily

frequency merges two distinct strands of the exchange rate prediction literature. The first strand

attempts to forecast exchange rates out-of-sample with ARIMA models (e.g., see Neely, Weller,

and Dittmar, 1997).[1] The second strand uses lower frequency Markov switching models to forecast exchange rates (Engel and Hamilton, 1990; Engel, 1994); this has enjoyed limited success.[2]

We view our work as complementary to that of Dewachter (2001) and Clarida, Sarno, Taylor, and Valente (2003), who use elements of both threads in very different ways. Dewachter (2001) constructs a Markov switching model that is able to produce simulated weekly data on which trend-following trading rules are successful. Clarida, Sarno, Taylor, and Valente (2003) generate an exchange rate forecasting model from a regime-switching vector error correction model (VECM) using weekly term structure data on forward exchange rates.

The risk-adjusted Markov trading rule returns exceed those of the ex ante best conventional technical trading rules considered for three of the four exchange rates. The differences are not statistically significant, but the Markov rules have two marginal benefits: 1) a portfolio of Markov and conventional technical rules has better risk-adjusted performance than either individually; and 2) Markov rules appear to do better on recent data, on which conventional trading rules are no longer successful. This is the best that can be expected; the noise in exchange rate returns makes it difficult to reject hypotheses of interest about trading rules.

Neither the returns to our trading rule nor the trading rule positions show a simple relationship with observable macro fundamentals across either sample period. But they may be

---

[1] Taylor (1994) reports some success with ARIMA models in the 1978-87 period, and we are able to extend those results to obtain some success in our out-of-sample period. These ARIMA rules are not as successful as the moving average or Markov rules studied here, however. LeBaron (1992) generated some trading rule profits with an ARIMA(1,0,1) fitted to maximize trading rule profits by simulated method of moments. This implies that nonlinearities are not strictly necessary to justify the profitability of moving average trading rules.

[2] Neely and Sarno (2002) review the literature on forecasting exchange rates with monetary fundamentals.

weakly related to higher moments of the data. Our model assumes that returns interact with a changing risk environment (higher moments) and the data are consistent with this proposition.

2.   Methodology

A. *The Markov switching models*

Let us first introduce some notation. The exchange rate at date t (USD per unit of foreign currency) is given by S_t, while r_t is the log of the deviation from uncovered interest parity, and the domestic (foreign) overnight interest rate is i_t (i_t^*).

$$r_{t+1} = \ln S_{t+1} - \ln S_t + \ln(1 + i_t^*) - \ln(1 + i_t) \tag{1}$$

The premise of trading strategies is that deviations from uncovered interest parity (UIP) are predictable. To create a rich structure for the expected deviation from UIP, we allow the conditional mean to be a function of three distinct Markov switching state variables. We assume a student-t error distribution with n_t degrees of freedom in the dependent variable r:

$$r_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim student-t(mean = 0, n_t, h_t), \qquad n_t > 2 . \tag{2}$$

The variance of such a student-t distribution is

$$\sigma_t^2 = h_t n_t /(n_t - 2). \tag{3}$$

The parameter h_t is a scale parameter for the variance such that $(r_t - \mu_t)/h_t^{1/2}$ is a standard student-t variable with n_t degrees of freedom. h_t switches between high and low states, according to the realization of a binary variable, $S1_t$, governed by the following first-order Markov process:

$$h_t = h_0 S1_t + h_1(1 - S1_t) \tag{4}$$

$$S1_t \in \{0,1\}, \qquad P(S1_t = 0 \mid S1_{t-1} = 0) = p_1, \qquad P(S1_t = 1 \mid S1_{t-1} = 1) = q_1.$$

Switching in h_t scales the variance up and down without affecting the thickness of the tails or leptokurtic shape of the conditional density, thus we refer to h as the dispersion parameter.

3

Similarly, we allow for switching in the degrees of freedom, n_t, as in Dueker (1997), which implies that the thickness of the tails of the conditional distribution varies across time. Thus, three- or four- standard deviation shocks can be statistically a near-impossibility in some time periods and a reasonable possibility in other periods. Because the kurtosis of a student-t variable, which equals $3(n-2)/(n-4)$, has a one-to-one relationship with the degrees of freedom, we refer to n as the kurtosis parameter. This kurtosis parameter is tied to a second binary variable, $S2_t$, that follows a Markov process such that

$$n_t = n_0 S2_t + n_1(1 - S2_t) \qquad (5)$$

$$S2_t \in \{0,1\}, \quad P(S2_t = 0 \mid S2_{t-1} = 0) = p_2, \qquad P(S2_t = 1 \mid S2_{t-1} = 1) = q_2.$$

The conditional mean, μ_t, is a function of the Markov state variables that govern switching in the dispersion and kurtosis. A third Markov switching binary variable, $S3_t$, provides yet another independent source of shifts in the expected return:

$$\mu_t = \mu_0 + \mu_1 S1_t + \mu_2 S2_t + \mu_3 S3_t \qquad (6)$$

$$S3_t \in \{0,1\}, \quad P(S3_t = 0 \mid S3_{t-1} = 0) = p_3, \quad P(S3_t = 1 \mid S3_{t-1} = 1) = q_3,$$

where μ_1 and μ_2 reflect how the dispersion and kurtosis affect the mean return. The three Markov states switch independently.[3]

Christoffersen and Diebold (2003) demonstrate that serial dependence in higher moments, such as the variance and kurtosis, affects the expected sign of returns in the presence of a non-zero unconditional mean return. This sign dependence creates predictability in the direction of returns. Our model does not directly exploit this effect; instead, it exploits dependence between conditional moments to better estimate the conditional mean.

---

[3] We also estimated a similar Markov model with 6 parameters in the mean, rather than 4. Results from the 2 models were very comparable.

For example, a rise in volatility can generate a change in conditional mean return through safe-haven effects. Higher volatility causes investors to seek safe-haven currencies, like the dollar. Decomposing volatility into both time-varying kurtosis and dispersion might improve our model's ability to detect the type of risk response associated with safe-haven effects.

In this vein, note that variation in the rate of information arrival affects the kurtosis of returns, whereas a shift in the marginal effect of a unit of information would affect the variance but not the kurtosis of returns. Thus, these features of the forecast distribution tell us different things, and our model allows them to covary with the expected return in distinct ways.

Although the model's expected return—conditional on the state—takes on only $(2^3=)$ 8 different values, because we do not observe the Markov state variables, we must infer the probability of the current state and hence the expected return. Therefore, the forecast conditional mean can take values from a continuum, not just a discrete number of values:

$$E[\mu_t \mid I_{t-1}] = \mu_0 + P(S1_t = 1 \mid I_{t-1})\mu_1 + P(S2_t = 1 \mid I_{t-1})\mu_2 + P(S3_t = 1 \mid I_{t-1})\mu_3. \qquad (7)$$

The log-likelihood function for this model is

$$\ln L(r_t \mid S1_t, S2_t, S3_t) = \ln(\Gamma(.5(n_t + 1))) - \ln(\Gamma(.5n_t)) - .5\ln(\pi n_t h_t)$$
$$- .5(n_t + 1)\ln\left(1 + \frac{(r_t - \mu_t)^2}{n_t h_t}\right) \qquad , \qquad (8)$$

where Γ is the gamma function and μ_t depends on the state variables, as in equation (6). Hamilton (1990) shows that the function to be maximized is the log of the expected likelihood or

$$\sum_{t=1}^{T}\ln\left(\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1}\text{Prob}(S1_t = i, S2_t = j, S3_t = k \mid I_{t-1})L(r_t \mid S1_t, S2_t, S3_t)\right). \qquad (9)$$

B. The trading rule from the Markov switching model

Although the model's parameters were chosen to maximize the log-likelihood function rather than returns, the model's trading performance is a more economically relevant measure of its fit

than any statistical measure.[4] This section describes how the model generates trading decisions.

A trading rule maps an information set (e.g., past exchange rates) to a binary variable, z_t, that takes the value +1 for a long position in foreign exchange at time t and –1 for a short position. Whereas traditional technical trading rules might use a moving average of prices, for example, to determine the trading signal, our likelihood-based approach can use the whole ex ante returns distribution to construct a trading signal.

A natural approach would be to have the trading rule go long if the expected deviation from uncovered interest parity, $E[\mu_t \mid I_{t-1}]$ from equation (7), is positive, and short otherwise. However, such a trading rule might generate large transactions costs by trading each time it predicts even a small non-zero return. We alleviate this problem in two ways. First, we use the expected return over eight days—rather than one—to make the trading decision. Second, we require that the expected return exceed a threshold, called a "filter," before we permit the rules to change position. Both techniques reduce trading frequency and accompanying transactions costs.

The period of eight days over which to calculate the expected return was chosen based on a representative half-life of the conditional mean of staying in the same state, using the transition probabilities estimated with in-sample data.[5] The average expected return over the eight-day

---

[4] Attempts to choose parameters to directly maximize in-sample returns proved unsuccessful as the return function is not continuously differentiable. Maximization methods—such as genetic algorithms—that work on non-differentiable functions might be helpful, but the success of more conventional methods defers that for the future.

[5] The average of the half-lives of the states varies across currencies and across the initial state. We chose an eight-day horizon as a representative value of the currency-specific averages. The implied half-lives were calculated only with transition probabilities estimated from in-sample data. Of course, the parametric model would allow us to choose currency-specific forecast horizons.

half-life is $\frac{1}{8}\sum_{k=1}^{8}E[\mu_{t+k}\,|\,I_t]$. We use this average expected return over eight days to avoid

cases where the expected return the next day is large in absolute value but is not expected to last

long before reversing. The eight-day average ensures that expected returns are persistent before

transactions costs are incurred.

A complete trading rule includes a pair of filter sizes, $f_1 < f_2$, that determine when one

actually changes position:

$$\text{If } z_{t-1} = +1, \qquad z_t = +1, \quad \text{if } \frac{1}{8}\sum_{k=1}^{8}E[\mu_{t+k}\,|\,I_t] \geq f_1$$
$$= -1, \quad \text{otherwise.}$$

$$\text{If } z_{t-1} = -1, \qquad z_t = -1, \quad \text{if } \frac{1}{8}\sum_{k=1}^{8}E[\mu_{t+k}\,|\,I_t] < f_2$$
$$= +1, \quad \text{otherwise.}$$

(10)

For example, the first two conditional equations above say that, if the rule has a long position at

t-1, it will only switch to a short position at t if the expected average exchange rate change from t

to $t + 8$ is less than the size of the filter, f_1. If the expected average change in the exchange rate

over the next 8 days is greater than or equal to f_1, the rule will maintain a long position.[6]

We chose filter sizes—given the parameters of the Markov switching model—to maximize

the excess return, net of transactions costs, in the in-sample (1974-81) period. Absolute filter

sizes ranged from about 0.2 to 6.25 basis points. The filter sizes were chosen under the

assumption of fairly high transactions costs of 10 basis points per trade. Such costs are about

---

[6] Note that the inertia in Markov-rule positions generated by the filter could require the Markov rule to stay in a

position in which it would be expected to lose (a little) money. It might seem wise to permit the rules to move to a

neutral position in such a situation. This intuition ignores the fact that such moves incur transactions costs.

twice the size that even a small trader could obtain today but probably reflect those faced by such

traders 20 years ago. Conservative transaction costs reduce the danger of overfitting.

C. Excess return calculation

The Markov switching rules switch between long and short positions in the foreign

currency. A long position in the foreign currency at date t means that the rule borrows dollars,

converts them to foreign currency at the closing rate for date t, and earns the foreign overnight

rate. A short position borrows in the foreign currency to invest in U.S. dollars. The excess

returns to long and short positions are closely approximated by $z_t r_{t+1}$, where z_t is the trading

indicator variable defined in equation (10), and r_{t+1} is the deviation from UIP, given by equation

(1). The cumulative excess return r for a trading rule giving signal z_t at time t over the period

from time zero to time T, conducting n trades, with transaction cost c, is as follows:

$$r = \sum_{t=1}^{T} z_t \, r_{t+1} - nc \, . \tag{11}$$

D. Economic performance measures: risk-adjusted excess returns

The excess return criterion tells us nothing about the trade-off between risk and return. This

information is absolutely necessary to evaluate the rule. The Sharpe ratio measures the expected

excess return per unit of risk for a zero-investment strategy (Campbell, Lo, and MacKinlay,

1997); it is the portfolio's mean annual excess return over the return's annual standard deviation.

But the Sharpe ratio only describes the univariate risk associated with a trading strategy.

Capital Asset Pricing Model (CAPM) theory tells us that the correlation of a portfolio's return

with that of the market should explain its excess return. Therefore we look at Jensen's (1968) α

and the CAPM β from the following regression:

$$z_t r_{t+1} - I(z_t \neq z_{t-1})c = \alpha + \beta[\ln(P_{t+1}/P_t) - \ln(1+i_t)] + \varepsilon_t, \tag{12}$$

where $z_t r_{t+1}$ is the signed return to the trading rule, $I(z_t \neq z_{t-1})$ takes the value one when a trade

is made, zero otherwise, and $[\ln(P_{t+1}/P_t) - \ln(1+i_t)]$ is the excess return to the market. Jensen's

(1968) α directly estimates the mean excess return that is uncorrelated with the market return. If

the intercept in (12)—α—is positive and significant, then the excess returns to the trading rule

cannot be explained by correlation with market returns. This study uses the Morgan Stanley

Capital International World Index (MSCI) and the S&P 500 to proxy for the market portfolio.

E. Statistical performance measures

To supplement the economic risk-adjusted return criterion, we also consider the performance

of the model's return predictions with respect to the mean-squared error (MSE), mean-absolute

error (MAE), and percentage of correct sign predictions at forecast horizons of 1, 5, 10, and 20

business days. The benchmark for comparison is the in-sample mean return.

3. The Data

The exchange rate data consist of noon (New York time) buying rates for the German mark,

euro, Japanese yen, British pound, and Swiss franc (DEM, EUR, JPY, GBP, and CHF) from the

H.10 Federal Reserve Statistical Release. Values for the EUR replace the DEM values after

1998, but we refer to the spliced series as the DEM series for simplicity. Exchange rates are

expressed as USD per unit of foreign exchange. The Bank for International Settlements (BIS)

provides daily interest rate data, collected at 9:00 am GMT (4:00 am, New York time).

To evaluate whether foreign exchange intervention or particular macroeconomic conditions

were associated with the trading rule returns or the smoothed probabilities of regime states, we

will use intervention data provided by central banks and monthly data from Haver Analytics on

industrial production, interest rates, stock market indices, M2, and employment from the five

countries: United States, Germany, Japan, Switzerland, and the United Kingdom.

Full summary statistics for the exchange rates are omitted here, but we have a few points to note: Two skewness statistics have inconsistent signs in the two subsample periods. But all foreign exchange excess return series are strongly leptokurtic over the whole sample and the degree of excess kurtosis has likely decreased from the 1970s to the 1980s and 1990s.

4.   Estimation results

A. *Estimation*

We estimate the Markov model using daily data from 1974 through 1981 (through 1982 for the yen/dollar rate). The in-sample periods were chosen to minimize the in-sample cumulative deviation from UIP, in order to guard against a rule that would tend to be disproportionately long or short. This reflects our prior belief that long-run expected deviations from UIP are probably close to zero for major currencies and that data that conform to this conviction are most likely to produce profitable rules. We did not use any information about the out-of-sample period to choose the in-sample period. Data from 1982 (1983 for the yen/dollar rate) through June 2005 are reserved to evaluate trading rule performance over a long out-of-sample period. Table 1 shows the log-likelihoods and parameter values from each exchange rate's best Markov model.

Each day, the model's predicted value for the exchange rate and a vector of filter sizes were used to map the data to a trading decision, using equation (10). The best filter sizes were chosen to maximize the in-sample excess return, given the estimated parameters.[7] Trading signals were then used to compute in- and out-of-sample trading return statistics.

Table 2 shows the trading rule statistics. The top rows show the filter sizes chosen on the basis of in-sample information. The rest of the table shows the number of observations, the

---

[7] The signs of the filters were not constrained.

annualized return, net of 10-basis-point transactions costs, in percentage terms, as well as the t-statistic for that annual return, mean trades per year, the percentage of business days the trading rule was long in the foreign currency, the Sharpe ratio and CAPM βs using the Morgan Stanley Capital International World Index and the S&P 500 as the market portfolios. The leftmost panel shows the in-sample results (1975-1981/82), the middle panel shows the full out-of-sample results (1982/83-2005:6) and the rightmost panel shows a 5-year subsample breakdown.

The absolute filter sizes ranged from about 0.21 to 6.25 basis points. (Expected returns are much less variable than actual returns.) The largest filters were associated with the CHF. The in-sample mean excess returns (leftmost panel of Table 2) were excellent but varied considerably among the 4 exchange rates, ranging from 10.30 to 16.80 percent per annum. The rules traded between 1.5 and 9.65 times per year in the in-sample period and were long between 56 and 71 percent of the time in the foreign currency.

B. Out-of-sample results

Of course, good in-sample performance is not very interesting for its own sake. The key issue is the out-of-sample risk-adjusted return of the rules. The middle panel of Table 2 shows that the out-of-sample excess returns are much lower than the in-sample returns but still good, ranging from 1.08 percent for the GBP to 7.54 percent per annum for the JPY. The average excess return over the four rates was 5.27 percent, a strong out-of-sample figure. The returns to the rules—except for the GBP—are statistically different from zero at any reasonable level with Newey-West standard errors with a lag order of 5. The rules trade 1.8 to 10.5 times per year in the out-of-sample period. The CHF rule is most often long in the foreign currency, taking a long position 74 percent of the time. The JPY rule is most often short in the foreign currency, taking a long position only 56.6 percent of the time.

11

Do the rules earn excess returns by taking on risk? It appears not. The risk-adjusted returns appear to be very attractive by any measure. The out-of-sample Sharpe ratios range from 0.11 (GBP) to 0.71 (JPY). The CAPM βs are close to zero, mostly statistically insignificant and negative as often as positive. There is no evidence that the rules take on excessive risk. Only the β for the GBP excess return versus the MSCI world index (β_1) is significantly positive.

Several authors have speculated that the returns to technical rules have declined in the last 10 or 15 years (Levich and Thomas, 1993; LeBaron, 2000; Okunev and White, 2003; Olson, 2004). Are the returns to these rules stable? Have they declined recently? We investigate this question in two ways: 1) a subsample breakdown of returns; 2) graphs of rolling Sharpe ratios.

The right-hand panel of Table 2 breaks down the mean returns by approximate 5-year subsamples. While the 1982-1991 period was very profitable for the rules, there is no clear evidence that the results are unstable or that the returns have disappeared. Mean returns over exchange rates are positive in all subsamples, including the most recent period, 2002-2005:6. Three of the four exchange rates have positive returns over the most recent subsample and the average over all four rates is a very respectable 5.11 percent. The GBP has the worst subsample performance with two subperiods (out of five) with negative returns. Of course, the high variability in exchange rate returns makes it difficult to see clear patterns.

Figure 1 shows the time series of the one-year moving average of Sharpe ratios to each rule. Although the out-of-sample Sharpe ratios are lower than the in-sample figures, there is no obvious break in the ratios. Figure 1 tells a similar tale to that in the subsample breakdown in Table 2: Sharpe ratios are usually positive, they showed some weakness in the 1990s but have rebounded lately. Wald tests for structural breaks in the mean return in the middle of the out-of-sample period (i.e., September 1993) fail to reject the null of no change at conventional levels for any of the four exchange rates. The DEM series produced the largest t statistic: 1.58. In contrast,

the same Wald tests for the MA rules clearly reject the nulls of equal means for the GBP and

CHF and marginally reject for the DEM (1.73). Full results are omitted for brevity.

The conclusion that one cannot reject stability in the Markov returns appears to be more

consistent with those of Okunev and White (2003) than Olson (2004). Olson (2004) found that

returns to trading rules that are reoptimized in rolling periods have been declining over time.

Perhaps the apparent contradiction should not come as a surprise; the structure and

reoptimization strategy of Olson's rules is very different than those examined here. Further,

Olson's sample period ended in 2000, prior to the most recent (and fairly profitable) subsample.

How do the Markov trading rules compare to traditional technical trading rules? To calculate

results from traditional technical rules, we compute signals from double moving average (MA)

and filter rules using the in-sample period of 1974-1981 (through 1982 for the JPY).[8] We

compute signals from all combinations of moving average (MA) rules with short moving

averages from 1 to 9 days and long moving averages from 10 to 150 days (in increments of 5

days). We permit the moving average rules to have "bands of inactivity," similar to those used

by the Markov trading rule (equation (10)). The MA band size and lag windows were chosen to

maximize in-sample returns. We compute signals from all combinations of filter rules with filter

sizes of 0.5 to 3 percentage points (in increments of 0.5) and go back 5 days to find extrema.

That is, we computed results for thousands of technical rules that are similar to those used in the

literature, such as in Neely (1997). We then found the best performing MA rule and filter rule

for each currency in the in-sample period and calculated out-of-sample trading rule statistics.[9]

Full results are omitted for brevity, but the technical rules do extremely well in the in-sample

---

[8] These "filter rules" have no direct relation with the filters used in the Markov trading rule, despite the name.

[9] The mean out-of-sample returns to the conventional technical rules were remarkably insensitive to changes in the

rule selection procedure such as elimination of the band, or using rolling or expanding in-sample periods.

period, as one might expect from the literature. The moving average rules' mean excess return was 14.5 percent over the four exchange rates and the filter rules had a 9.2 percent mean return. The MA and filter rules performance deteriorate in the out-of-sample period, but is still strong—especially that of the MA rules. The filter rules' annual returns range from –3.1 percent to 3.6 percent per annum and those of the MA rules from 3.41 percent to 5.94 percent. The average annual return for the better performing MA rules is 4.32 percent. These out-of-sample returns are somewhat lower than those found by Sweeney (1986) and Neely (1997) using shorter data samples. But the overall MA returns are higher than those found by Olson (2004) who reports returns from dynamically reoptimized technical rules declining from 3 percent in the 1970s and 1980s to about zero in the 1990s. Indeed, the MA rules have a small negative mean return after 1991. Because the MA rules outperform the filter rules, this paper will concentrate on comparing the Markov rules to the better performing MA rules. The Sharpe ratios and CAPM βs also show little evidence that the technical rules' excess returns are compensation for bearing risk. The Sharpe ratios are very good—those of the MA rules range from 0.32 to 0.56, averaging 0.4—and the CAPM βs are small and insignificant. In short, these results are consistent with those found previously for technical trading rules in foreign exchange markets (e.g., Neely, 1997). Technical trading rules make excess returns that cannot be explained by transactions costs or the usual risk adjustments.

As good as the traditional technical rules' performance is, however, the mean annual returns and risk-adjusted returns for the Markov trading rules are about 95 basis points better, on average, across the four exchange rates. This advantage in net returns stems from the fact that the Markov rules trade much less often than the MA rules. The MA rules trade 20 or 30 times a year, in contrast to the 2 to 11 times a year for the Markov rules. The only case in which the double MA rules outperform the Markov rules—by 2.3 percentage points—is that of the GBP.

And the average Sharpe ratio for the Markov rules is 0.48, compared to 0.40 for the MA rules.

We are reluctant to claim, however, that the Markov rules really improve on the MA rules. Tests of differences in mean returns are generally insignificant—though the JPY test has a p-value of 0.065. (Results omitted for brevity.) The Markov rules do appear to have two important marginal benefits for investors, however: 1) A portfolio rule actually outperforms either the Markov rules or the MA rules on a risk-adjusted basis; and 2) The Markov rules (and the portfolio rules) are much more profitable than the MA rules in the most recent period.

The fact that the Markov and technical trading rules have very different trading frequencies indicates that the two types of rules are finding different sorts of trends and that combining them could improve over the univariate risk-return performance of either type, separately. [10] Table 3 shows the results from a portfolio rule whose signals are made up of equal shares of Markov signals and MA rule signals.[11] The returns and trades are the average of Markov and MA returns, of course, but the Sharpe ratios—the annual mean return over the annual standard deviation of returns—show the marginal contribution of the Markov rules.[12] The Sharpe ratios for the portfolio rule are better than those of the MA rules in 3 of 4 cases and nearly as good in the GBP case. (MA rule results are omitted for brevity.) The average Sharpe ratio for the portfolio rules is 0.54, versus 0.48 for the Markov rules and 0.40 for the MA rules. The improvement in Sharpe ratios is surely related to the imperfect correlation in the rule signals from the two models. The Markov and MA models produce the same signals 74, 70, 66, and 55 percent of the time for the DEM, JPY, CHF and GBP, respectively. In addition to higher Sharpe

---

[10] It is certainly possible, however, that different sorts of technical rules could closely approximate the Markov rules.

[11] A mean-variance optimal portfolio rule produced similar results to the equally weighted rule.

[12] A rule with a higher Sharpe ratio is superior to a second rule with a higher return because leverage can be increased on the former—with a commensurate increase in risk—to provide a superior return per unit of risk.

ratios, the right-hand panel of Table 3 shows only 4 subsamples (of 20) of negative returns for the portfolio rules, compared to 5 negative subsamples for the Markov rules and 6 periods for the MA rules. (MA subsample results are omitted for brevity.) That is, the portfolio rules have the most robustly positive returns over subsamples; the Markov rules are second best. This advantage of the Markov rules is shown most clearly in the most recent subsample (2002-2005:6), where the Markov rules had a 5.11 percent return versus the –3.34 percent return for the MA rules. Risk-averse technical traders would have been much better off combining the Markov rules and the MA rules for the last 23 years, rather than using the MA rules alone.

To investigate whether the success of the Markov rules hinges on promptly entering the market when a trend is spotted, we recalculated the rules' performance with lagged trading signals. The results—omitted for brevity—show only a modest (0.8 percentage point) diminution in returns to the Markov rules from the contemporaneous signal results depicted in Table 2. The MA rules suffer a larger (2 percentage point) loss in profitability. The Markov rules exploit longer-term trends, not just returns immediately after a trade.

C. Statistical criteria

Finally, we compare MSE and MAE for the Markov switching model to the naive constant return model at forecast horizons of 1, 5, 10, and 20 business days. The ratios less than one in the upper panel of Table 4 show that the Markov model consistently forecasts better than the naive model, in-sample. But this forecasting advantage fails to consistently carry over to the out-of-sample period (bottom panel). In the out-of-sample period, the constant return and Markov models provide nearly identical forecasts by the MSE and MAE criteria. The Markov model probably does relatively better on the MAE criterion because the MSE is more influenced by extreme errors. The fact that the Markov model predicts the sign of the return correctly supports this view. The "% Right" is above 50 for every exchange rate, at all horizons, in both

16

subsamples, and significantly above 50 percent in 9 of 16 out-of-sample cases. As one might

expect from the return calculations, the CHF and GBP models have relatively weaker "% Right"

out-of-sample forecasting performance. In summary, the Markov model successfully predicts

the sign of returns, despite its mixed performance on the MSE and MAE criteria.

5.   What Creates Trading Rule Returns?

A. *Discussion of the source of trading rule returns*

    Perhaps the most important unsettled issue in studies of technical analysis is the source of the

returns. Technical analysts credit psychology for the success of their methods: Asset traders will

tend to react the same way when confronted by the same conditions: Past price patterns will

predict future price patterns (Neely, 1997). Economists have tried to explain the success of

technical analysis with models of asymmetric information/sequential trading, behavioral finance

induced biases, and intervention by government authorities (LeBaron, 1999; Szakmary and

Mathur, 1997; Saacke, 2002; and Sapp, 2004).

        Although the full results are omitted for brevity, there appears to be no consistently

significant relationship between the trading rule profits and/or probabilities of the state variables

and the international differences in growth rates of macro-variables. Further, there is no

evidence that intervention causes trading rule returns. Neely (2002) used high-frequency returns

and intervention to show that the timing and direction of trading are inconsistent with the idea

that central bank intervention generates technical trading rule profits.

        Although we cannot find a relation between trading rule returns/signals/states and

observable macro variables, one might be able to find a relation between the excess returns and

higher conditional moments of the return distribution.

17

B. Excess returns and higher moments

To investigate the source of the excess returns, we examine whether the interaction between returns and their conditional higher moments is important through both statistical tests and trading rule returns. That is, does restricting the contribution of dispersion and kurtosis degrade the fit of the model and the trading rule returns?

To test this, we maintain switching in the dispersion and kurtosis (so that the transition probabilities remain identified), but shut down the mean switching related to dispersion and kurtosis. We can calculate likelihood ratio (LR) test statistics for these restrictions on the effect of the dispersion and kurtosis state variables on the mean:

$$dispersion: \mu_1 = 0 \qquad \text{and} \qquad kurtosis: \mu_2 = 0$$

For each of the four currencies, the LR test statistic is significant at or near 5 percent for at least one of the two restrictions. The smallest of each currency's p-values are: DEM 0.057 for mean switching tied only to dispersion, h; JPY 0.039 for mean switching tied only to kurtosis, n; CHF 0.012 for mean switching tied only to h; GBP 0.060 for mean switching tied only to n.

In addition to the statistical evidence of fit, we calculate trading returns for Markov trading rules that prevent mean switching tied to dispersion and kurtosis. Restricting the mean of the Markov model from using higher moments reduces overall mean annual returns by 1.5 percentage points, to 3.77 percent. The diminution in mean out-of-sample returns ranged from 34 (CHF) to 272 (DEM) basis points for the four currencies. Sharpe ratios were 14 basis points lower, on average, for the restricted model. The t statistics for differences in mean returns between models with 4 mean parameters versus 2 were 1.87, 0.85, 0.15, and 0.60 for the four exchange rates, using Newey-West standard errors. The t statistic for the difference in the portfolio returns (over the 4 exchange rates) was 1.4. The number of trades and the percentage long varied over currencies but was fairly comparable across models, overall. The signals from

18

the restricted model agreed with the signals from the unrestricted model 70 to 85 percent of the time. In summary, the data are consistent with the hypothesis that using higher moments increases returns but the evidence is not conclusive.

6.   Conclusions

This paper has used Markov switching models to create ex ante trading rules in the foreign exchange market. Markov models generate statistically and economically significant out-of-sample returns that are 95 basis points larger, on average, than those of conventional technical trading rules, and these returns appear to be fairly stable over time. The Markov rules provide at least two marginal benefits over conventional MA rules. An equally weighted portfolio rule of the Markov and MA rules provides a better risk-return trade-off than either alone. In addition, the Markov rules are strongly superior to the MA rules on the most recent data, in which the MA rules' profitability seems to have disappeared.

The Markov switching models deliver strong out-of-sample portfolio returns, although they fail to outpredict a naive, constant-return benchmark by MSE and MAE criteria. While the mean returns have diminished after 1991, tests reject structural breaks in Markov mean returns, which are still positive in every subsample, including the period from 2002 to 2005:6. Thus, Markov rule returns have been more stable than those of the conventional MA rules.

The ability of the Markov trading rules to identify trends in exchange rates might be linked to their use of information about higher moments. The fact that in-sample LR tests always preferred linking either the distribution's dispersion (scale of the variance) or kurtosis to the mean return supports this contention. Restricting the mean of the Markov model from using higher moments reduces overall mean annual out-of-sample returns by 1.5 percentage points and Sharpe ratios by 14 basis points. This suggests, but does not prove, that higher moments belong

19

in the expectations of the Markov trading rule. The technical trading literature has not previously exploited higher moments in constructing rules.

The use of econometric methodology, rather than technical rules, to make trading decisions has at least two potential advantages. First, one can generate the entire multi-period distribution of exchange rate returns, enabling the risk-averse investor to better assess the risk-adjusted expected returns. A second potential advantage of an econometric methodology is that the stability of the model structure—rather than the return moments—can be assessed in real time, enabling traders to change their trading rules with the structure of the data-generating process. This paper did not explore those advantages.

References

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Chang, P.H.K., Osler, C.L., 1999. Mythological madness: technical analysis and the irrationality of exchange-rate forecasts. Economic Journal 109, 636-661.

Christoffersen, P.F., Diebold, F.X., 2003. Financial asset returns, direction-of-change forecasting and volatility dynamics, unpublished.

Clarida, R.H., Sarno, L., Taylor, M.P., Valente, G., 2003. The out-of-sample success of term structure models as exchange rate predictors: a step beyond. Journal of International Economics 60, 61-83.

Dueker, M., 1997. Markov switching in GARCH processes and mean-reverting stock market volatility. Journal of Business and Economic Statistics 15, 26-34 .

Dewachter, H., 2001. Can Markov switching models replicate chartist profits in the foreign exchange market? Journal of International Money and Finance 20, 25-41.

Engel, C., 1994. Can the Markov switching model forecast exchange rates? Journal of International Economics 36, 151-165.

Engel, C., Hamilton, J.D., 1990. Long swings in the dollar: are they in the data and do markets know it? American Economic Review 80, 689-713.

Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39-70.

Jensen, M.C., 1968. The performance of mutual funds in the period 1945-1964 (in Problems in selection of security portfolios). Journal of Finance 23, 389-416.

LeBaron, B., 1992. Do moving average trading rules results imply nonlinearities in foreign exchange markets? University of Wisconsin-Madison Social Systems Research Institute Working paper 9222.

LeBaron, B., 1999. Technical trading rule profitability and foreign exchange intervention. Journal of International Economics 49, 125-143.

LeBaron, B., 2000. Technical trading profitability in the 1990s. Unpublished manuscript, Brandeis University.

Levich, R., Thomas, L., 1993. The significance of technical trading rule profits in the foreign exchange market: a bootstrap approach. Journal of International Money and Finance 12, 451-474.

Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: do they fit out of sample? Journal of International Economics 14, 3-24.

Neely, C. J., 1997. Technical analysis in the foreign exchange market: a layman's guide. Federal Reserve Bank of St. Louis Review 79(5), 23-38.

Neely, C. J., 2002. The temporal pattern of trading rule returns and central bank intervention: Intervention does not generate technical trading rule profits. Journal of International Economics 58, 211-32.

Neely, C. J., Sarno, L., 2002. How well do monetary fundamentals forecast exchange rates? Federal Reserve Bank of St. Louis Review 84(5), 51-74.

Neely, C. J., Weller, P., Dittmar, R., 1997. Is technical analysis in the foreign exchange market profitable? A genetic programming approach. Journal of Financial and Quantitative Analysis 32, 405-426.

Okunev, J., White, D., 2003. Do momentum-based strategies still work in foreign currency markets? Journal of Financial and Quantitative Analysis 38, 425-447.

Olson, D., 2004. Have trading rule profits in the currency market declined over time? Journal of Banking and Finance 28, 85-105.

Saacke, P., 2002. Technical analysis and the effectiveness of central bank intervention. Journal of International Money and Finance 21, 459-479.

Sapp, S., 2004. Are all central bank interventions created equal? An empirical investigation. Journal of Banking and Finance 28, 443–474.

Sweeney R.J., 1986. Beating the foreign exchange market. Journal of Finance 41, 163-182.

Szakmary, A.C., Mathur, I., 1997. Central bank intervention and trading rule profits in foreign exchange markets. Journal of International Money and Finance 16, 513-535.

Taylor, S. J., 1994. Trading futures using a channel rule: A study of the predictive power of technical analysis with currency examples. Journal of Futures Markets 14, 216-235.

Table 1: Parameter values

		DEM	JPY	CHF	GBP
S1	μ_1	0.002	0.082	0.058	-0.030
	h_0	0.044	0.012	0.062	0.005
	h_1	0.313	0.258	0.419	0.154
	p_1	0.970	0.973	0.990	0.955
	q_1	0.977	0.987	0.992	0.988
S2	μ_2	0.049	-0.006	0.131	0.012
	$1/n_0$	0.005	0.005	0.096	0.006
	$1/n_1$	0.301	0.359	0.370	0.465
	p_2	0.991	0.969	0.985	0.978
	q_2	0.995	0.989	0.974	0.986
S3	μ_3	0.212	0.243	0.257	-0.283
	p_3	0.989	0.989	0.984	0.920
	q_3	0.998	0.995	0.997	0.995
	μ_0	-0.221	-0.228	-0.287	0.300
	LL	-1394.315	-1358.414	-1852.541	-1057.298

Notes: The table shows the estimated parameters and the log-likelihoods from the model described in equation (8), estimated over the in-sample period: 1974 through 1981 (through 1982 for the USD/JPY rate). S1, S2, and S3 are the Markov state variables.

Table 2: Markov trading rule statistics

	DEM	JPY	CHF	GBP
Filter 1	-0.0072	-0.0128	-0.0625	-0.0021
Filter 2	-0.0023	0.0211	0.0374	0.0369

	In-Sample: 1974 - 1981				Out-of-Sample: 1982 - 2005:6				Subsample return breakdown				
	DEM	JPY	CHF	GBP	DEM	JPY	CHF	GBP		DEM	JPY	CHF	GBP
Return	10.30	16.80	12.04	11.85	6.42	7.54	6.06	1.08	1982-1986	9.58	12.78	10.88	1.30
t-stat	3.05	5.22	2.80	3.81	2.93	3.36	2.50	0.53	1987-1991	13.20	7.01	9.99	7.26
trades	9.65	5.23	1.50	1.75	10.52	7.78	2.00	1.83	1992-1996	-0.47	11.14	4.89	-6.98
% long	69.15	60.20	70.94	55.97	61.11	56.60	73.91	72.10	1997-2001	2.53	6.65	-2.84	-0.99
Sharpe	1.08	1.74	0.99	1.35	0.61	0.71	0.52	0.11	2002-2005:6	7.64	-1.66	7.99	6.47
Beta 1	NA	NA	NA	NA	-0.02	0.01	0.02	0.17					
SE 1	NA	NA	NA	NA	(0.05)	(0.05)	(0.05)	(0.04)					
Beta 2	-0.08	0.03	-0.07	0.02	-0.09	-0.02	-0.08	0.02					
SE 2	(0.07)	(0.07)	(0.08)	(0.07)	(0.04)	(0.05)	(0.05)	(0.04)					

Notes: The rows in each panel show the filter sizes (see equation (10)), the number of observations, the annualized return, net of transactions costs, in percentage terms, as well as the t statistic for that annual return, the mean number of trades per year, the percentage of business days the trading rule was long in the foreign currency, the Sharpe ratio and CAPM βs using the Morgan Stanley Capital International World Index and the S&P 500 as the market portfolios. β_1 is the CAPM beta of the trading rule portfolio with the MSCI world index and β_2 is the analogous statistic for the S&P 500. The leftmost panel shows the in-sample results (1974-1981/82), the middle panel shows the full out-of-sample results (1982/83- June 2005) and the rightmost panel shows a subsample breakdowns of the mean return. The initial subsample excludes 1982 for the JPY.

Table 3: Tests of portfolio rules using equally weighted shares of the Markov trading rules and the best, ex ante moving average rules

Out-of-Sample: 1982 - 2005:6					5-year Subsample Returns				
	DEM	JPY	CHF	GBP		DEM	JPY	CHF	GBP
Return	6.18	5.73	5.04	2.23	1982-1986	11.18	10.61	10.02	5.22
t-stat	3.29	2.86	2.57	1.42	1987-1991	12.86	7.99	9.95	11.16
trades	16.36	19.34	11.63	14.57	1992-1996	-1.35	6.94	2.88	-5.20
% long	42.36	38.18	44.30	38.77	1997-2001	3.33	5.07	0.21	-2.23
Sharpe	0.69	0.64	0.53	0.31	2002-2005:6	4.36	-3.96	0.91	2.22

Notes: The left-hand panel displays out-of-sample results for an equally weighted portfolio rule using Markov and MA signals. The right-hand panel shows returns for the portfolio rule over subsamples. See the Notes to Table 2 for row headers.

Table 4: Forecast statistics for the Markov switching models

| | In-Sample | | | |
	DEM	JPY	CHF	GBP
MSE 1	0.992	0.990	0.996	0.999
MSE 5	0.978	0.947	0.992	0.993
MSE 10	0.968	0.952	0.994	0.989
MSE 20	0.968	0.955	0.998	0.991
MAE 1	0.995	0.993	0.995	0.997
MAE 5	0.992	0.985	0.992	0.995
MAE 10	0.988	0.988	0.994	0.993
MAE 20	0.986	0.993	0.997	0.992
%Right 1	53.37	55.06	53.32	54.67
%Right 5	53.37	54.92	54.22	54.12
%Right 10	52.02	54.84	52.32	53.77
%Right 20	51.42	53.06	50.87	53.37

| | Out-of-Sample | | | |
	DEM	JPY	CHF	GBP
MSE 1	1.002	1.001	1.002	1.001
MSE 5	1.001	1.002	1.000	1.001
MSE 10	1.002	1.002	1.001	1.000
MSE 20	1.002	1.000	1.001	1.000
MAE 1	1.000	1.004	1.001	1.001
MAE 5	0.999	1.002	0.999	1.000
MAE 10	0.999	1.001	1.000	0.999
MAE 20	1.000	0.999	1.001	0.999
%Right 1	51.62	51.40	50.52	50.76
%Right 5	51.94	51.42	51.71	51.42
%Right 10	51.64	51.40	50.28	51.06
%Right 20	50.52	51.35	50.31	50.98

Notes: The top panel shows the number of observations and the ratios of in-sample MSE and MAE for the Markov switching model to the naive constant return model at forecast horizons of 1, 5, 10, and 20 business days. Ratios less than one indicate that the Markov model forecasts exchange rate returns better than the naive model. The rows labeled "%Right 1" to "%Right 20" show the percentage of predictions of the return with the correct sign at the same horizons. The bottom panel shows the out-of-sample results. Asymptotic standard errors for the %Right statistics are $\sqrt{p(1-p)/T} = 1.117\%$ for the top panel and 0.65% for the bottom panel.

Figure 1: One-year rolling Sharpe ratios
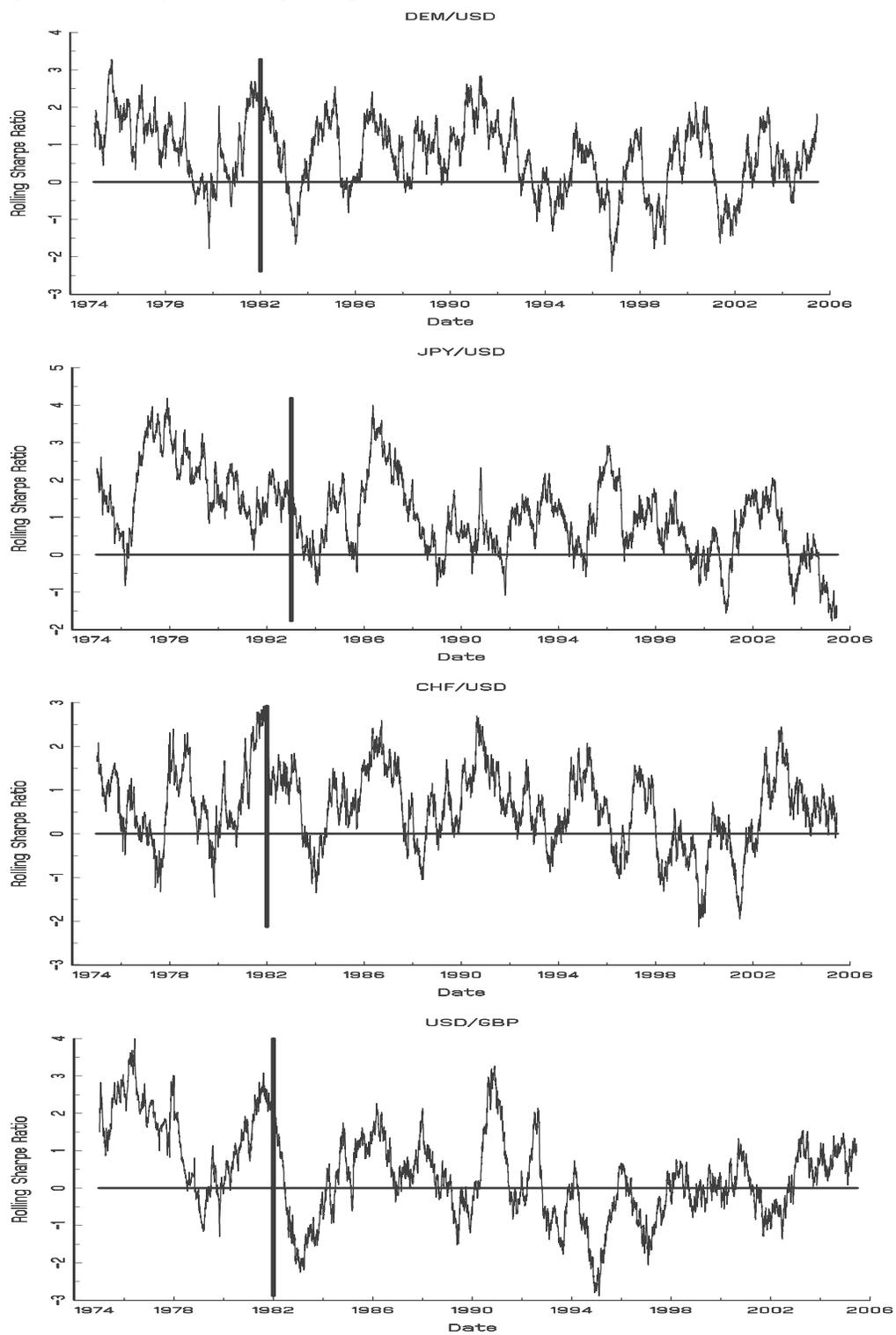
Notes: One-year rolling Sharpe ratios over the whole sample. Vertical lines depict the break

between in-sample and out-of-sample periods. Horizontal lines denote zero.