Technical Trading Profitability in Foreign Exchange Markets in the 1990's

Blake LeBaron *

September 2002

Preliminary and incomplete

Abstract

This paper presents evidence on the changes in the performance of technical trading rules in foreign exchange markets during the 1990's. Previously reported good performance for earlier time periods is no longer as strong. This dramatic shift is used as an experiment to explore whether its cause could be related to data snooping, or deeper economic issues.

*Graduate School of International Economics and Finance, Brandeis University, 415 South Street, Mailstop 021, Waltham, MA 02453 - 2728, blebaron@brandeis.edu, www.brandeis.edu/~blebaron. The author is also a faculty research fellow at the National Bureau of Economic Research, and an external faculty member at the Santa Fe Institute.

1 Introduction

Recent evidence has shown that simple technical trading rules have the ability to predict movements in foreign exchange prices, and to create successful dynamic trading strategies. This paper looks at the performance of these strategies over the decade of the 1990's and finds earlier claims to their performance to be somewhat exaggerated. These results suggest several interesting puzzles that will be explored using some new bootstrap related time series tools.

The natural first possibility is that foreign exchange markets have changed. If this is the case, several important economic issues are opened in trying to explain this change. Could it be the result of reduced transaction costs, foreign exchange intervention, the internet, or maybe better general knowledge of these types of rules? Before one is ready to consider all these possibilities it is important to rule out one other possibility, data snooping. The ever present problem in economic time series analysis is certainly an issue here. Proponents for data snooping explanations would suggest that these results are nothing more than an artifact of fitting an optimal trading rule to one sample (pre 1990), and then looking at it out of sample (post 1990). The dramatic change is nothing more than an indication of these biases. Although, the results have remained strong over several early subperiod studies, and most of the papers in this area have tried to only use commonly traded rules, the snooping question can never really be escaped. Several different bootstrapping and simulation methodologies will be implemented to try to answer these questions.

The second section will give an overview of the data. The third section looks at the basic trading rule returns. The fourth assesses the regime change. The fifth section looks at the hypothesis of a stationary model as the generating process, and the final section concludes.

2 Data

The data used in this paper are daily foreign exchange rates from June 6, 1973 through Aug 28, 2002. They are collected by the Federal Reserve Bank of New York, and are noon (Eastern Standard Time) rates. Interest rates are overnight rates from the Bank of International Settlements which are measured at the European market close.[1]

Most of this paper will be concerned with returns after adjustments for interest differentials. Taking S_t

[1]They are out of synch with the exchange rates by several hours. This may be a problem for detailed technical analysis evaluation, but experiments have shown that the exact timing of interest rates is not important. For Japan the BIS rates start in 1983. Earlier rates use daily quotes of monthly Japanese short rates from 73-77, and daily quotes of 1 month euro rates from 77 to 83. The goal was to construct the longest possible consistent series of FX rates, with some preference to getting the exchange rate prices correct.

as the nominal exchange rate in U.S. \$/ FX the adjustments are made as follows.

$$s_t = \log(S_t)$$

$$y_{t+1} = (s_{t+1} - s_t + r_t^* - r_t)$$

where r_t^* is the foreign interest rate, and r_t is the U.S. interest rate. y_t is therefore the interest adjusted return. In a risk neutral world with uncovered interest parity this should have expected value zero, and be uncorrelated with time $t-1$ information. Also, if covered interest parity holds it is the return to forward speculation given by

$$y_{t+1} = (s_{t+1} - f_t).$$

where f_t is the one period ahead forward rate. These are all the returns to a zero cost strategy of borrowing in one currency and investing it in another.

Table 1 presents summary statistics for excess returns for the three currencies. They are shown for the British Pound (BP), German Mark (DM), and Japanese Yen (JY). In each case the series is broken down into the entire sample along with the pre and post 1990 subsamples. The table shows nothing unusual about the series in terms of relative high frequency financial time series. They all appear near zero mean, with relatively large kurtosis. The subsample breakdown shows no remarkable differences across the two subsamples either.

Table 1: *Summary Statistics: Excess Returns*

	Mean (percentage)	Std.	Skewness	Kurtosis
BP	-0.0008	0.61	-0.12	6.91
BP (73-89)	-0.0049	0.63	-0.05	7.55
BP (90-2002)	0.0047	0.57	0.57	5.62
DM	-0.0012	0.66	-0.04	6.22
DM (73-89)	-0.0014	0.66	-0.10	7.55
DM (90-2002)	-0.0046	0.67	0.04	4.52
JY	0.0036	0.66	0.37	8.63
JY (73-89)	0.0071	0.61	0.16	10.21
JY (73-2002)	-0.0011	0.72	0.54	7.17

All returns are daily returns adjusted for interest differentials. The entire sample contains 7325 daily observations. The first subsample contains 4156 observations, the second subsample contains 3169 observations, and the final subsample contains 1408 observations.

3 Trading Rule Returns

Most of the tests in this paper are concerned in some way with the performance of moving average technical trading rules. The remarkable thing about them is not that really fancy rules work, but that the most basic and simple technical trading strategies appear to offer enhanced performance. This paper concentrates on only very simple moving average rules. These are formed by deciding on a buy or sell signal given the level of today's price relative to moving average of past prices.

$$m_t = \frac{1}{M} \sum_{i=0}^{M-1} S_{t-i}$$

If $S_t > m_t$ this is a buy period, and the signal, $z_t = +1$. If $S_t \leq m_t$ then it is a sell period and the signal $z_t = -1$. The final object of interest is the dynamic strategy given by

$$E(z_t y_{t+1})$$

This trading rule moment forms the basic object of interest for most of the tests in this paper. It is estimated using the time averages of the strategy. Also, in most cases M is set to 150 days. This has been found to be a generally reliable strategy, and it is also a very common one which has been used for many years, and many markets. This makes it less vulnerable to datasnooping arguments, even though some of these will be tested directly in this paper.

Table 2 presents a summary of the trading rule moments for a 150 day moving average rule. For the entire sample, all three currencies yield a significant trading rule moment ranging from 0.01 to 0.03 percent per day. The next to last column estimates the Sharpe ratio which is simply the trading rule moment, which is an excess return, divided by its standard deviation. This is multiplied by $\sqrt{250}$ giving an annualized Sharpe ratio. A common benchmark for this simple risk adjusted measure is that an aggregate equity portfolio buy and hold strategy gives a value ranging from about 0.3 to 0.4. The dynamic FX strategies all perform well relative to this simple benchmark. The final column estimates a standard error on the Sharpe ratio by bootstrapping the dynamic strategy moment 1000 times. This table displays the different subsamples as well. For the BP and DM these display a very different picture of the results. For the BP series the returns actually turn negative, but insignificant. For the DM the returns drop down to 0.016 percent in the last decade with a Sharpe ratio reduction of 0.58 to 0.37. Also, the t-test on the excess return for the DM is no longer significant with a value of only 1.31. The JY shows similar reductions, but the last decade still looks

Table 2: *Dynamic Strategy Returns*

	Mean (percentage)	T-test (mean=0)	Sharpe Ratio	std(Sharpe)
BP	0.013	1.83	0.34	0.19
BP (73-89)	0.027	2.73	0.68	0.25
BP (90-02)	-0.004	-0.45	-0.13	0.29
BP (97-02)	-0.028	-2.15	-0.91	0.44
DM	0.020	2.58	0.48	0.19
DM (73-89)	0.023	2.30	0.58	0.25
DM (90-02)	0.016	1.31	0.37	0.27
DM (97-02)	0.016	1.00	0.42	0.42
JY	0.031	3.92	0.73	0.18
JY (73-89)	0.036	3.80	0.95	0.26
JY (90-02)	0.023	1.80	0.51	0.28
JY (97-02)	0.017	0.82	0.34	0.42

Table 3: *Trading Rule Moment Differences*

	Break Date	Difference 10^{-4}	T-test
BP	30-Dec-1989	3.31	2.20
DM	30-Dec-1989	0.72	0.47
JY	30-Dec-1989	1.20	0.76

interesting. The reduction continues with a very large reduction over the past 5 year period where it is no longer signficant.[2]

Figure 1 presents a rolling 5 year t-test for the 150 day moving average dynamic strategy on each of the 3 currencies. The date on the x-axis corresponds to a 5 year window moving forward from that date. The change point appears to be close to the arbitrarily chosen decade change point of 1990. The next section will further explore these dramatic changes.

4 Regime Shifts

This section explores the possibility of a regime shift in a more precise fashion. Specifically, the series of trading rule returns given by x_t is assumed to be independent over time for the purposes of resampling to generate new time series of pseudo trading rule returns using the IID bootstrap. Comparisons are made across both a fixed break point at 1990, and a moving break point determined by the maximum difference between the trading rule moments.

Table 3 presents results for a simple test of equality in the trading rule moments across a fixed break point at 1990. The t-test tests equality of the means over the two subsamples. It is interesting that only for

[2] Judging a drop in the t-test is difficult since the sample size is reduced dramatically in the last subsample.

Table 4: *Trading Rule Moment Differences: Variable Break*

	Break Date	Mean 10^{-4}	P-value
BP	July-1991	4.97	0.001
DM	January-1992	2.06	0.281
JY	September-1986	2.15	0.240

the the BP is this test significant. The actual difference from the original series

Table 4 adds the feature of choosing the break point. In this case, the break point that maximizes the difference in the original series is chosen by searching from 1985 through 1995. It is necessary to stop early to keep a large enough part of the sample in the later half. The results of this table are very close to those of the previous table and for the BP and DM series, the breakpoint is very close to the simple 1990 calendar breakpoint. However, for the DM and JY the differences are not significant when compared with a bootstrap of other maximized breakpoint choices. The JY remains a series with no clear break, and the it identifies a point well before 1990.

Another method for testing whether the last several years have been unusual is subsampling. In this procedure five or ten year long periods are chosen at random from the first subsample. The length depends on the length of the later subsample. This is done 1000 times, and the fraction of randomly drawn subperiod with trading rule moments larger than the recent past period is recorded. It is reported in table 5. This table gives the mean annualized return over the last 10 and 5 year periods. These are then compared a simulated distribution of 10 or 5 year returns drawn at random. Subsample mean reports the mean of this distribution, and subsample std. the standard deviation. P-value is the fraction of these simulations that are larger than the the recent period subsamples. For the 10 year periods all three series show that the last period was unusual. For the 5 year case, only the JY and BP series look different.

The subsampling approach is easy to understand, and is appealing because it replicates all the dependencies present in the original series. However, as the samples are lengthened it gets a relatively poorer monte-carlo distribution since there are not all that many long periods to draw at random without much overlap. Some of this may be happening here, but it is also possible that this test is doing a better job of including dependencies in the series that the other tests miss.

5 Testing Stationarity

The previous section provides evidence that something has changed in at least two of the three foreign exchange series. This section provides several tests of a stationary null hypotheses for the series. Table 6

Table 5: *Trading Rule Comparisons: Subsamples*

Comparison Range	Currency	Annualized Return	Subsample Mean	Subsample Std.	P-value
1990-2002	BP	-1.16	8.27	0.61	0.00
	DM	3.86	7.36	0.47	0.00
	JY	5.78	11.50	0.36	0.00
1997-2002	BP	-6.73	6.94	2.96	0.00
	DM	4.19	6.61	2.88	0.224
	JY	4.39	9.44	3.25	0.043

Table 6: *IID Bootstrap*

Rule	Series	Rule	Difference	73-89	90-2002
Fixed	BP	150	0.003	0.000	0.806
	DM	150	0.152	0.009	0.155
	JY	150	0.143	0.000	0.048
Variable	BP	10	0.005	0.000	0.499
	DM	25	0.028	0.000	0.106
	JY	100	0.281	0.000	0.011

tests a constant random walk model across the entire sample. This is done by taking the log first differences of the foreign exchange price series, and sampling them with replacement. These scrabled differences are used to build new representative random walk series. The results in table 6 present the fraction of 1000 simulations generating a trading rule return as large as the actual data. Results of this simulated p-value are given for the two subsamples in the last two columns of the table. The first column examines the difference between the trading rule returns in the first and second subsamples and reports the fraction of simulations generating a difference as large as the actual series. The two parts of the table represent two different styles of testing. The first takes the 150 day moving average rule as given and stays with this through all tests. The second part of the table attempts to simulate a form of data snooping, by estimating the best rule on the first half of the series and using this rule for all tests. Rules are chosen from a set of $(10, 25, 50, 100, 150, 200, 250)$ day moving averages. The intention of this experiment is to answer questions about whether the differences in results across subsamples could be driven purely by snooping that took place in the first subsample alone.

For the first subsample, the IID bootstrap shows little ability to replicate the results. This is to be as expected since previous results have displayed similar findings with the null hypothesis of an independent random walk strongly rejected. The results begin to vary a little in the second subsample. The BP shows no evidence for deviations from IID in the second subsample representing the general weakening of the trend predictability in this period. Similar weaker results hold for the DM, but the JY continues to show a similar signficant pattern to the early subsample. These results are all reflected in the differences. For the fixed

Table 7: *Trend Model Parameters*

Series	Rule	ρ	σ_u^2/σ_e^2
BP	150	0.61	0.067
BP	10	0.63	0.062
DM	150	0.89	0.007
DM	25	0.94	0.004
JY	150	0.94	0.004
JY	100	0.96	0.003

rules the differences are only significant for the BP, but for the variable rules both the BP and DM represent a significant difference. The results need to be viewed cautiously since the underlying null model, the IID random walk is rejected in the data. This will be addressed as dependent FX models are explored in the next tables.

Foreign exchange series display both significant returns for trend following rules and very low autocorrelation in the first differenced series. To accomplish this a candidate model with a trend plus noise is used. It is given by,

$$r_t = x_t + e_t \tag{1}$$

$$x_t = \rho x_{t-1} + u_t \tag{2}$$

where r_t is the return, or log difference of the exchange rate levels at time t. u_t and e_t are both IID normal with variances of σ_u^2 and σ_e^2 respectively.

This equation is estimated using a crude method of movements estimator. For a given value of ρ, the estimated variance of r_t, and the first order autocovariance of r_t can be used to get values of σ_u^2 and σ_e^2, the variances of the two noise processes. The computer sweeps through values of ρ, and using the correponding variances numerically, estimates the trading rule return. ρ is chosen to minimize the absolute difference between the trading rule return in the simulation and the actual data. Table 7 gives the estimated parameter values for ρ and the ratio of the two variances. This ratio is a kind of signal to noise ratio indicator on the process.

The trend model is now used as a model for the entire series. Table 8 presents the results from a simulations of this model replacing the IID bootstrap simulations from the previous test. For the 150 day moving average rule the trend model misses significantly for the BP series on both subsamples, coming in too low on the first half and too high on the second. This inability to fit the two halves is demonstrated by the small p-value for the difference between the two samples. Results for the fixed 150 day moving average

Table 8: *Trend Sim Monte-Carlo*

Rule	Series	Rule	Difference	73-89	90-2002
Fixed	BP	150	0.010	0.054	0.975
	DM	150	0.189	0.280	0.753
	JY	150	0.169	0.177	0.685
Variable	BP	10	0.005	0.026	0.984
	DM	25	0.038	0.127	0.913
	JY	100	0.247	0.462	0.765

for the DM and JY series give a different story with results unable to reject the constant trend model as a possible explanation for the series dynamics in either subsample. When the variable MA rule is used the results do not change for the BP or JY, but the DM changes dramatically. The p-value for the difference now indicates that it is unlikely for the single trend model to generate as large a difference as that seen in the data. Also, the results for the subsamples are now farther out at the extremes.

This last table turns to a more nonparametric approach to handling the problem of replicating the dependence in the the series, the stationary bootstrap. The stationary bootstrap samples a series with replacement as in the standard bootstrap, but to simulate dependence in the series, it samples varying length blocks. The block sizes are determined randomly. When a point is drawn at random from the original series, the sampling procedure chooses the next occuring times series point with probability $(1 - \lambda)$, or a completely new time location with probability λ. This has the effect of building random length blocks in the resampled series, and replicating some of the dependence in the original time series.

For these simulations λ was fixed at 0.002 giving an average block size of 500. Results are given in 9. The results for the fixed moving average rule again show a strong difference between the subsamples for the BP, and the corresponding high and low values for the first and second subsamples respectively. The results are similar, but weaker for the DM and JY series. Turning to the adjustable rules the results again agree with the last table. The dramatic change is that for the optimal 25 day rule, the DM series shows a very strong subsample difference along with the appropriate extremes in the two subsamples. The first subsample data level is too large to be captured by blocks drawn from the entire sample, and the second is too low. This is consistent again with some kind of sample break. The JY series again shows little indication for signficant change.

Table 9: *Stationary Sim Bootstrap*

Rule	Series	Rule	Difference	73-89	90-2002
Fixed	BP	150	0.015	0.063	0.940
	DM	150	0.105	0.141	0.757
	JY	150	0.100	0.132	0.764
Variable	BP	10	0.007	0.050	0.952
	DM	25	0.012	0.054	0.957
	JY	100	0.157	0.590	0.680

6 Conclusions

The dramatic change in the behavior of dynamic technical trading strategies in the BP and DM series presents a very intriguing opportunity. This is both a potentially interesting economic event, and an opportunity to assess our ability to sort out data snooping biases from important regime shifts.

This paper has concentrated on the issues related to the econometrics of the problem. The results are generally supportive that there has been a change in regime in these two series. However, they should be viewed with some caution for two reasons. First, in many cases they are only marginally significant and hold for only two foreign exchange series. Tests with other series are clearly necessary. Second, the performance of the stationary bootstrap in small samples may not be very good in this situation. It is not clear whether this very convenient technique can nonparametrically assess data snooping biases for dynamic trading strategies.[3]

If the changes are real, then several different causes remain a possibility. First, LeBaron (1999) relates profitability to foreign exchange interventions. Schwartz (2000) notes that these interventions have dropped off for most currencies in the later 1990's. It is also possible that efficient markets are playing a role, and that traders are finally figuring these strategies out. Finally, it is possible that transactions costs may be falling over the period, allowing traders to better arbitrage, and again trade away these features. These difficult questions cannot be answered in this context.

One final conclusion is that no matter how you look at the data, any trader thinking about using these dynamic strategies should be very cautious. The evidence is weak enough to keep most anyone from jumping in. In a longer term sense it may be that these rules are profitable only over very long horizons, but can go through long periods in which they lose money. These periods shake out most of the users, and prepare the market for future trending periods since no one is paying attention. This also remains an interesting question both for the future of foreign exchange markets, and a kind of implicit measure of just how irrational they may have been in the past.

[3] It should be noted that Sullivan, Timmerman & White (1999) have much longer time series at their disposal.

References

LeBaron, B. (1992), Do moving average trading rule results imply nonlinearities in foreign exchange markets?, Technical report, University of Wisconsin - Madison, Madison, Wisconsin.

LeBaron, B. (1999), 'Technical trading rule profitability and foreign exchange intervention', *Journal of International Economics* **49**, 125–143.

Politis, D. & Romano, J. (1994), 'The stationary bootstrap', *Journal of the American Statistical Association* **89**, 1303–1313.

Schwartz, A. J. (2000), The rise and fall of foreign exchange market intervention, Technical Report 7751, National Bureau of Economic Research, Cambridge, MA.

Sullivan, R., Timmerman, A. & White, H. (1999), 'Data-snooping, technical trading rule performance and the bootstrap', *Journal of Finance* **54**, 1647–1691.

Taylor, S. J. (1980), 'Conjectured models for trends in financial prices, tests and forecasts', *Journal of the Royal Statistical Society A* **143**, 338–362.

Figure 1: *Rolling Trading Rule T-test*