EFDW
7/19

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

Blake LeBaron *

July 2000
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.

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 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 first section will give an overview of the data. The second section looks at the basic trading rule returns. The third assesses the regime change. The fourth section looks at the possible impact of data snooping. The fifth section assesses the performance of the stationary bootstrap, and the final section concludes.

2 Data

The data used in this paper are weekly foreign exchange rates from June 6, 1973 through May 27, 1998. They are all measured at the close of the London markets on Wednesdays or the following day on holidays. Offshore interest rates are used to adjust the foreign exchange returns for interest differentials. From 1979 to the present these are taken from 1 week eurorates from Natwest Bank. For the earlier time periods it was necessary to use other interest rate series as proxies for the 1 week eurorates. For the BP, DM, and U.S. Dollar, the rates are 1 month eurorates (sampled daily). For September 1977- December 1978 the Yen also uses 1 month eurorates. From June 73 through August 1977 the Japanese interest rates are Bank of Japan short rates (sampled monthly).[1]

Most of this paper will be concerned with returns after adjustments for interest differentials. Taking S_t 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. There is not even any hint of a regime change here.

Table 1: *Summary Statistics: Excess Returns*

	Mean	Std.	Skewness	Kurtosis
BP	0.018	1.42	-0.38	5.98
BP (73-89)	-0.003	1.41	0.09	5.14
BP (90-98)	0.059	1.42	-1.17	7.74
DM	-0.008	1.46	0.00	4.49
DM (73-89)	-0.016	1.44	0.19	3.92
DM (90-98)	0.005	1.51	-0.32	5.42
JY	0.003	1.41	0.33	5.22
JY (73-89)	0.017	1.38	0.38	5.48
JY (73-98)	-0.031	1.48	0.26	4.74

[1]During the periods after 1979 I have shown the results are relatively insensitive to the interest rate series, so it is unlikely they will have much of an effect here, (LeBaron 1999).

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 30 weeks. 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 30 week moving average rule. For the entire sample all three currencies yield a significant trading rule moment of about 0.10 percent for the BP and DM, and 0.18 percent for the JY. 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 $\sqrt{52}$ 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.

The story gets more interesting when one considers the period after 1990. For both the BP and DM the trading rule moments drop off dramatically along with the Sharpe ratio. The T-test for the mean indicates that the trading rule moment is not significantly different from zero in either case. Further, the Sharpe ratios for these two currencies are no longer very large, and are actually not even significantly different from zero.

Table 2: *Dynamic Strategy Returns*

	Mean (percentage)	T-test (mean=0)	Sharpe Ratio	std(Sharpe)
BP	0.109	2.78	0.56	0.20
BP (73-89)	0.149	3.02	0.76	0.25
BP (90-98)	0.050	0.86	0.25	0.29
DM	0.098	2.41	0.49	0.21
DM (73-89)	0.127	2.59	0.65	0.25
DM (90-98)	0.039	0.54	0.19	0.35
JY	0.178	4.48	0.91	0.20
JY (73-89)	0.187	3.89	0.97	0.25
JY (90-98)	0.155	2.19	0.76	0.34

Table 3: *Trading Rule Moment Differences*

	Break Date	Mean (% Difference)	P-value
BP	30-Dec-1989	0.110	0.088
DM	30-Dec-1989	0.087	0.153
JY	30-Dec-1989	0.032	0.342

As an interesting contrast, the results for the JY appear not to change very much across the subsamples.

Figure 1 presents a rolling 4 year t-test for the 30 day moving average dynamic strategy on each of the 3 currencies. The date on the x-axis corresponds to a 4 year window moving forward from that date. The results corroborate the table indicating a possible change for the BP and DM series, but little evidence for change in the JY. 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 for the DM and BP series.

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 bootstrap analyzing returns across a fixed break point at 1990. The actual difference from the original series is compared with scrambled series that assume returns are drawn from the same distribution both pre and post 1990. The table reports simulated p-values, or the fraction of bootstrap runs generating values as large as those from the original series. The results show moderately significant

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

	Break Date	Mean (% Difference)	P-value
BP	26-Jun-1991	0.227	0.020
DM	11-Dec-1991	0.178	0.080
JY	20-Aug-1986	0.119	0.212

differences for both the BP and DM series, and no significant difference in the JY. This is consistent with the simpler results given in the previous section.

Table 4 adds the impact of choosing the break point. In this case, the break point that maximizes the difference in the original series is chosen by searching from 1990 through March 11, 1991. This last date marks the point 3/4 through the entire sample. It is necessary to stop early to keep a large enough part of the sample in the later half. It is clear from the table that this search boosts the difference, and for both the DM and BP series the simulated P-value is further reduced. The JY remains a series with no clear break. Future tests will concentrate on only the BP and DM.

5 Data snooping and the stationary bootstrap

At first glance this may appear to be an obvious indication that something dramatic has changed in these series. However, the evidence should be viewed with some trepidation. There is an issue of data snooping the technical trading rule that is actually used. In all cases the 30 week moving average has been used to align with previous results and also because it was a commonly used rule by technical trading practitioners. It was not snooped in my previous papers, but there are still two lingering questions: it had already been snooped by practitioners and I had been exposed to this information when commencing studies in the late 1980's. In some way I snooped a little in that if the 30 week MA had not worked I would probably have tabled the earlier studies.

The data snooping bias story would say that the dramatic breaks are coming from the fact that the 30 week MA rule was tuned by someone to the first half of the sample, and it therefore no longer works well in the later part. The difference comes purely from snooping and not from an actual regime break. This would seem like a difficult hypothesis to tests, but recently the tools are becoming available to do this. The paper Sullivan, Timmerman & White (1999) proposes a method for estimating data snooping biases. This paper follows that in spirit if not in exact methodology.

In order to estimate a data snooping bias one actually needs two things. First, a new clean replication of

Table 5: *Stationary Bootstrap Tuning*

Series	λ	Mean $x10^4$	Std $x10^4$
BP	0.50	1.78	3.76
	0.75	3.77	4.09
	0.95	7.95	4.04
	0.99	9.06	3.91
	0.995	8.92	4.15
Actual		10.90	
DM	0.50	2.24	4.03
	0.75	4.23	4.16
	0.95	9.34	3.78
	0.99	12.1	3.43
	0.995	12.1	3.43
Actual		9.80	

the time series process. Second, an exact algorithmic approach to the snooping process. In other words one needs to know just which rules were snooped over. Obviously, in practice both of these are difficult. This paper follows Sullivan et al. (1999) in using the stationary bootstrap developed by Politis & Romano (1994). It also compares this with a simulated time series model calibrated to properties of the foreign exchange series.

The stationary bootstrap is a dependent sampling procedure designed to replicate arbitrary dependence in a time series. Instead of resampling a series independently blocks of variable length are sampled. The procedure is quite simple. When a new point is required one first goes to random point in the original series, call it τ. Now when the proceeding point at $t+1$ is needed one first draws a uniform random number on $u = [0, 1]$. If $u < \lambda$, where λ is a parameter in the stationary bootstrap, then the next point is taken from $\tau + 1$. If $u > \lambda$ then a completely new point in the original series is chosen. This gives a sequence of contiguous blocks of random length. The only major problem is the choice of λ. Table 5 analyzes this for the FX series that are being used here. It is probably a good idea to balance the choice of λ between replicating a feature of interest, and the length of the data. Too large a value of λ will give the mostly the original series back. The table shows results for several values of λ in terms of means and standard deviations on the 30 week technical rule.[2] In each case it is clear that only the values of $0.95, 0.99, 0.995$ are reasonable ones for the stationary bootstrap.

Table 6 shows the results for the BP, and the stationary bootstrap along with three values of λ. For each simulated series a search is conducted over the range of [5,50] week moving average rules to find the

[2]It is important to note that the bootstrap is being tuned on a potentially snooped value. This may be a problem, but it is difficult to figure out how to correct this.

Table 6: *BP Snooping Tests*

Method	λ	Mean $x10^4$	Std. $x10^4$	% >Actual Fixed break	% >Actual Variable break
Constant Break	0.95	3.23	7.90	0.168	0.007
	0.99	2.88	9.12	0.191	0.016
	0.995	3.15	9.90	0.215	0.030
Varying Break	0.95	8.05	7.11		0.034
	0.99	8.39	7.85		0.057
	0.995	8.40	8.41		0.071

Table 7: *DM Snooping Tests*

Method	λ	Mean $x10^4$	Std. $x10^4$	% >Actual Fixed break	% >Actual Max Variable break
Constant Break	0.95	1.86	7.81	0.206	0.020
	0.99	2.82	7.78	0.238	0.032
	0.995	2.88	8.22	0.240	0.043
Varying Break	0.95	6.83	6.67		0.076
	0.99	7.51	7.30		0.103
	0.995	7.42	7.19		0.105

one which maximizes the return over the period before 1990. In the first panel of the table a fixed break at 1990 is used, and compared to the fixed break difference from the original series. A variable break is used in the second panel which replicates the search from 1990 through 3/4 through the sample as done previously. The table reports both the mean and standard deviation for each difference along with the fraction of 1000 simulation runs which generate values as large as those in the actual series. The column labeled with Fixed Break uses the fixed break difference from the original series as the critical value, and the column labeled variable break uses the variable break point value. The table gives a more conservative report on the BP in terms of a break point when using the fixed point. However, for the variable break the evidence still points to a break with a relatively small fraction of runs generating a value greater than the actual data.

Table 7 presents a similar story for the DM. There appears to be little indication of a break when using the fixed point, but the variable break point yields simulated fractions that are also nearly as small as for the BP. According to these tables the impact of data snooping on the break point evidence would be small.

The previous results have relied on a new technology, the stationary bootstrap, which remains somewhat untested in small samples. To give some perspective on this, a more traditional test will now be used. A simple stochastic trend model is used to simulate the long horizon trends in foreign exchange markets and replicate the trading rule profits while maintaining low autocorrelations. Stochastic trend models have been presented as possible explanations for trends and trading rules by Taylor (1980). They also have been fit to

Table 8: *Trend Model Tests*

Series	Trading Return	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
BP	0.140	0.049	0.044	0.044	0.040	0.039
	(0.701)	(0.246)	(0.879)	(0.841)	(0.192)	(0.127)
DM	0.109	0.070	0.012	0.014	0.067	0.073
	(0.604)	(0.081)	(0.103)	(0.916)	(0.580)	(0.789)

foreign exchange series using simulated method of moments estimators in LeBaron (1992).

The time series models used here are given by

$$x_{t+1} = \mu_t + \eta_t \tag{1}$$

$$\mu_{t+1} = p\mu_t + e_t \tag{2}$$

where both η_t and e_t are independent Gaussian noise. Instead of trying to exactly fit this model to the data as in LeBaron (1992) it will be calibrated to the actual series using some rough estimates. The hope it just to provide a generically reasonable model for the foreign exchange series which can give us a comparison with the stationary bootstrap. Calibration entails setting p high enough to get the persistence large enough so that the technical rules work, and to keep the autocorrelations in the series low. A given value of p will be guessed, and then other parameters will be chosen using analytic moments to match the variance of the original series, and to minimize

$$\sum_{j=1}^{5}(\rho_j - \hat{\rho}_j)^2 \tag{3}$$

where ρ_j is the autocorrelation at lag j for the trend model, and $\hat{\rho}_j$ is the estimated autocorrelation from the original series. This locks down all the parameters in the trend model. Table 8 shows the properties of a fitted trend model. In both cases p is fixed at 0.95 with the other parameters fit to the individual foreign exchange series. The numbers in parenthesis are the fraction of 1000 simulations generating values greater than the original series. With an exception of a few autocorrelations in the DM series the fit appears pretty good.[3]

Table 9 repeats the results of tables 7 and 6 using the simulated time series as opposed to the stationary bootstrap. The simulations are run both using a constant and varying break point. For the BP series the results with the stand in time series model do not differ from those with the stationary bootstrap. The simulated p-value for the varying break case is about 6 percent. The results change slightly for the DM

[3]It should be noted here that one obvious feature that this model is missing is the persistence in volatility. This may be contributing to some of the problems in yielding a good fit for the DM.

Table 9: *Time Series Snooping Tests*

Series	Method	Mean	Std.	% >Actual	% >Actual Max
BP	Constant Break	3.56	10.24	0.227	0.026
BP	Varying Break	8.71	8.14		0.063
DM	Constant Break	4.37	10.21	0.339	0.096
DM	Varying Break	9.18	8.33		0.184

Table 10: *Bias Estimation*

Test	Mean	Mean Bias	% > True	Std	Skew	Kurtosis
True	13.61	-	-	0.54	-0.262	2.67
MC	16.21	2.61	0.70	5.02	0.069	3.04

series. In this case the varying break case generates little evidence for a significant break with a simulated p-value of 0.184.

This opens a question of why there is such a difference. One possibility is that the time series model is simply not well specified which is always going to be the case in the parametric time series world. Another possible problem is that the sample sizes are too small for the stationary bootstrap. To address this latter question an experiment is performed to test the reliability of the stationary bootstrap on a known time series process, and data snooping procedure. The process will be the one used for the BP series. This process will be used to generate representative time series of length 1000. For these the stationary bootstrap will be used to estimate the biased value of the trading rule moment given that the rule is chosen in the set $[5, 50]$ over the sample. The procedure begins by finding the "true" trading rule moment by simulating a very long, 100,000, time series, and again searching for an optimal rule in [5,50]. A simulation of 100 of these long series shows that the "true" value is estimated to be $13.61x10^{-4}$ which is reported in table 10. The standard deviation shows that this is precisely estimated. In the second part of the experiment a picture of the bias in searching over rules is given by actually simulating multiple shorter length (1000) time series, and finding the optimal rule. The value of $16.2e10^{-4}$ shows a clear upward bias over the true value. Also, the table reports in column 4 that 70 percent of these runs were greater than the true value. There is a clearly some in sample bias here.

Finally, table 11 attempts to assess the bias using a stationary bootstrap. Bias is estimated by taking the original sample as the population, and the bootstrapped version as a sample. For each bootstrapped sample the difference is found for,

$$E(x^*_{m^*} - x_{m^*}) \tag{4}$$

where $x^*_{m^*}$ is the value of the technical trading moment using the moving average length adjusted to that

Table 11: *Bias Simulations*

λ	Mean $x10^4$	Std. $x10^4$	Fraction < True
0.95	0.44	0.11	1.00
0.99	1.35	1.41	1.00

sample, and x_{m^*} is the technical trading moment using the same moving average length, but estimated on the original sample. This bias is estimated over 1000 bootstraps, and is simulated for 100 draws of the original series. The means reported in table 11 show the bootstrap is clearly underestimating the true bias. This is further emphasized by the fact that none of the simulations generated a bias as large as for the actual data.

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.[4]

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

[4]It should be noted that Sullivan et al. (1999) have much longer time series at their disposal.

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.

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*