

THE EFFICIENCY OF THE FORWARD EXCHANGE MARKET:  
A CONDITIONAL NONPARAMETRIC TEST OF FORECASTING ABILITY

Roy D. Henriksson  
Donald R. Lessard\*

SSM-WP- 1337-82

July 1982

\*The Authors are Assistant Professor of Finance, Graduate School of Business, University of California, Berkeley, and Associate Professor of Management, M.I.T. Sloan School of Management. They are grateful to Arnold Barnett, Fischer Black, Ilene Jacobs, and Robert Merton for helpful comments.

Working papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the authors. For that reason working papers may not be reproduced or distributed without the written consent of the author.

## Abstract

This paper presents a conditional, nonparametric test of the information efficiency of the forward exchange market. This test is superior to parametric tests since it does not require restrictive assumptions regarding the distribution of changes in foreign exchange rates and does not involve a joint test of a particular model of the equilibrium relationship between forward and expected future spot rates. It is superior to unconditional nonparametric tests since, although they also do not require restrictive assumptions regarding the distribution of exchange rate changes, they do incorporate the implicit hypothesis that the forward rate is an unbiased estimator of future spot rates. Further, it allows many more independent observations within a given test interval than either parametric or unconditional nonparametric tests, thus increasing its power. Using weekly forecasts from two services, some evidence of forecasting ability is found. More importantly, the unconditional results show that conditional tests lead to both type I and type II errors regarding forecasting ability.

## I. INTRODUCTION

The volatility of foreign exchange rates and the perception of market inefficiency have been a cause of concern among investors, corporate managers, and economic policy makers. Government intervention in the foreign exchange markets has been considered as a response to these concerns and the potential impact of such intervention has been the issue of much debate. A central issue to this debate has been the efficiency of the market in incorporating new information into prices. Does intervention improve the efficiency of the market or does the interference by governments in the market or only lead to greater profit opportunities for speculators?

This paper focuses on the relationship between the forward rate and the corresponding expected future spot rate and presents a methodology for testing the efficiency of the forward exchange market. Previous tests of this relationship have been critically dependent on the assumed model of exchange rate determination used in the tests. As all tests of market efficiency are really tests of the joint hypothesis of market efficiency and the validity of the assumptions necessary to conduct the tests, uncertainty regarding the true model of exchange rate determination casts doubt upon the results of previous tests. Even if the tests reject their null hypothesis, it is not possible to determine if market efficiency is being rejected or just the validity of the assumptions used in the test.

Most previous tests of this type also have suffered from the additional problem that they have been limited to small number of independent observations because of the limited history of floating exchange rates and the fact that the shortest duration for a standard forward contract is one month. This has been exacerbated by the fact that most tests have focused on three or six month forwards as more reliable and relevant time series.

Because of the limited number of independent observations and the volatility of exchange rates, even if the true model of forward rate determination is known, it is quite likely that evidence of market inefficiency would not have been detected even if it did exist.

Most of the problems of past tests of forecasting ability can be avoided using the nonparametric techniques first applied to financial forecasts by Henriksson and Merton [1981]. These tests are derived from the basic model of forecasting ability developed by Merton [1981] where the forecaster predicts direction, but not magnitude. Using this framework, the tests evaluate forecasting ability without requiring any assumptions regarding the model of forward rate determination or the distribution of future spot or forward rates. By definition, superior forecasting ability must be based on information that is not reflected in prices. Therefore, the existence of superior forecasting ability is a violation of market efficiency. The nonparametric nature of the tests also makes it possible to increase substantially the number of independent observations, therefore greatly increasing the power of the tests of forecasting ability.

The main goal of this paper is to present a methodology for testing market efficiency through the evaluation of forecasting ability. The methodology is demonstrated by evaluating the forecasts of two foreign exchange advisory services.

Problems with previous tests of forward market efficiency are discussed in Section II. These problems are primarily the result of the lack of a precise model for the determination of the forward rate, but also result from the restrictive distributional assumptions required for parametric tests. The statistical techniques used by Henriksson and Merton are described in Section III. A way to substantially increase the number of

independent observations which requires virtually no additional assumptions is presented in Section IV, along with a test of the hypothesis that the probability of the future spot rate exceeding the forward rate is equal to 50 percent. The hypothesis appears to be violated for a number of currencies. In Section V, the results of the nonparametric tests of forecasting ability, requiring no assumptions about the model of forward rate determination, are presented for two foreign exchange advisory services. Some evidence of forecasting ability is found. Comparisons are drawn with unconditional, nonparametric tests which are shown to lead to both Type I and Type II errors regarding forecasting ability.

## II. PROBLEMS WITH PAST TESTS

Tests of the efficiency of foreign exchange markets can be divided into three groups: tests of the interest rate parity theory (IRPT), tests of spot market efficiency, and tests of forward market efficiency. In this paper, the focus is on the efficiency of the market for forward foreign exchange. A test of market efficiency is presented that examines forecasting ability with respect to the relationship between the forward rate and the corresponding future spot rate. The importance of this relationship should be clear as forward contracts provide a mechanism for eliminating undercertainty resulting from exchange rate exposure.

Two recent papers, by Hansen and Hodrick [1980] and Bilson [1981], have provided valuable insights into the relationship between the forward rate and the expected future spot rate. Levich [1979] and Kohlhagen [1978] both provide surveys of previous tests that focus on this relationship. However, all of these previous studies are not true tests of market efficiency

because of uncertainty regarding the true model of forward rate determination. Instead, they are descriptive of the time series relationship between the forward rate and the corresponding realized spot rate.

Most previous tests have been parametric in nature as they have depended on the values of  $[s(t+n) - f(t,n)]$  where  $s(t+n)$  is the actual realized spot rate of exchange between two currencies at time  $(t+n)$  and  $f(t,n)$  is the forward rate at time  $t$  for the spot rate at time  $(t+n)$ . All tests based on returns or differences in forward rates and the corresponding realized spot rate require knowledge of the model of forward rate determination to correctly account for risk. This includes evaluations of forecasting ability where the evaluation is based on the magnitude of the difference or the return from an investment strategy based on the forecasts.

The simplest assumption is that the forward rate is an unbiased estimate of the expected future spot rate, implying that there is no risk premium embedded in the forward rate. The validity of this assumption is suspect, however, as it is certainly possible to construct a reasonable model of forward rate determination based on market efficiency and rational expectations on the part of investors where the forward rate is not equal to the expected future spot rate because of risk aversion or the costs of trading and information.

Much work, both theoretical and empirical, has been done on this subject. Papers by Grauer, Litzenberger, and Stehle [1976], Kouri [1977], and Fama and Farber [1979] show that the forward rate can include a risk premium because of the uncertainty of the relative inflation rates. All three of these papers assume a one-period model where investors maximize a utility function which is an increasing function of expected terminal real

wealth and a decreasing function of the variance of terminal wealth. Fama and Farber also show that only if purchasing power parity holds and if all individuals have identical consumption baskets will exchange rate uncertainty be irrelevant for portfolio decisions, and, hence, the forward rate will be an unbiased estimator of future spot rate.

Using the same assumptions, Frankel [1979] shows that the existence of outside assets<sup>1</sup> in the economy will cause the forward rate to include a risk premium, even if the real rate of return for the economy is independent of the exchange rate.

Stulz [1981] examines the implications for the exchange rate when the assumptions of a one-period equilibrium and identical consumption baskets are relaxed. He shows that when individuals in different countries have different consumption baskets, the uncertainty of the real exchange rate can result in the forward rate including a risk premium that is a function of the level of net domestic foreign investment. This is true even if the nominal exchange rate is not correlated with real returns in the economy or if there are no outside assets.

When the assumption of a one-period equilibrium is relaxed, Stulz shows that the risk premium embedded in the forward rate may also reflect the correlation of changes in the exchange rate with intertemporal sources of risk, such as changes in the investment opportunity set. The intertemporal models of asset valuation derived by Merton [1973] and Breeden [1978] provide a framework for evaluating the intertemporal implications of exchange rate determination. Using a model similar to that of Breeden, Stulz shows that the risk premium embedded in the forward rate is an increasing function of the correlation of changes in the domestic exchange rate with changes in aggregate real world consumption.

Recent empirical studies by Hansen and Hodrick [1980] and Bilson [1981] both find evidence that the hypothesis that the forward rate is an unbiased estimate of the expected future spot rate can be rejected for a number of major currencies (relative to the U.S. dollar). This can be the result of either market inefficiencies or the existence of a risk premium. Unfortunately, using the parametric techniques of the studies, it is not possible to distinguish between these two possibilities.

Because of the lack of a precise theory of forward rate determination and the fact that expectations are not observable, it is necessary to use historical data for the forward rate and the corresponding realized spot rate to estimate the relationship. Further, most methods used to estimate the risk premium require that the relationship be stationary. However, even if one assumes that the foreign exchange market is efficient, it is quite difficult to estimate the risk premium embedded in forward rates. Because of the volatility of exchange rates, a long period of time would be required for estimation as the risk premium is certainly small relative to the standard deviation of the exchange rate. Therefore, estimation is critically dependent on the assumption of stationarity. Bilson [1981] discusses the problem of estimating a risk premium included in forward rates and Merton [1980] provides an excellent description of the problem of estimating the mean in the context of the risk premium embedded in the expected return on the market portfolio in the United States.

In addition, since it is necessary to assume market efficiency to estimate the relationship between the forward rate and the expected future spot rate, it is virtually impossible to test for market efficiency when knowledge of the relationship is necessary.

Even if the risk premium included in forward rates is known, parametric tests suffer from the additional problem that it is necessary to know the characteristics of the distribution of exchange rates for hypothesis testing. The usual assumption is that the exchange rate follows a normal distribution. However, a number of recent papers<sup>2</sup> have questioned the validity of this assumption.

Because of these problems with parametric tests, researchers have increasingly turned toward nonparametric tests of market efficiency.<sup>3</sup> Such tests typically involve counting the percentage of periods that a forecaster is exact, with 50 percent as the critical point. Such tests require stationarity in the relationship between the forecasts and the exchange rate and the assumption that forecasters are predicting direction, but not magnitude. With respect to foreign exchange forecasting, the forecaster predicts whether or not the forward rate will be greater than the corresponding future spot rate, but not by how much.

Nonparametric tests based on the unconditional probability of a correct forecast assume that the probability of a correct forecast is independent of the magnitude of the difference between the forward and the actual future spot rate. Because of this, unconditional tests are really tests of the joint hypothesis of no forecasting ability and the assumption that the probability of each of the two possible outcomes (either  $s(t+n) > f(t,n)$  or  $f(t,n) \leq s(t+n)$ ) occurring is 50 percent. In foreign exchange evaluation, this requires that the probability of the realized future spot rate exceeding the forward rate is 50 percent<sup>4</sup> for symmetric distributions. This implies that the forward rate is an unbiased estimate of the expected future spot rate, a hypothesis that has been rejected for many currencies.<sup>5</sup> Therefore, nonparametric tests of the unconditional probability of a correct forecast are suspect as tests of market efficiency.

### III. CONDITIONAL, NONPARAMETRIC TECHNIQUES FOR TESTING FORECASTING ABILITY

The uncertainty regarding the true model of forward rate determination has made it impossible to test the efficiency of the market for forward exchange contracts using parametric techniques. In addition, unless the probability of the realized future spot rate exceeding the forward rate equals 50 percent, a questionable assumption, nonparametric tests of forecasting ability that focus on the unconditional probability of a correct forecast will not be tests of market efficiency. Given the state-of-the-art for models of forward rate determination, what is necessary to test the efficiency of forward rates is a technique that requires no assumptions about the relationship between the forward rate and the expected future spot rate. This can be accomplished through the use of a nonparametric test based on the conditional probabilities of a correct forecast, conditional upon whether or not  $s(t+n) > f(t,n)$ .

Merton [1981] developed a framework for evaluating forecasting ability that does not require knowledge of the distribution of the forecasted variable or any particular model of security valuation. In the foreign exchange market, the forecaster predicts the relationship between the forward rate at time  $t$  for the spot rate at time  $(t+n)$ ,  $f(t,n)$ , and the actual spot rate at time  $(t+n)$ ,  $s(t+n)$ . The forecaster is assumed to simply predict (or only has the ability to predict) that the forward rate will exceed the future spot rate [i.e.,  $f(t,n) > s(t+n)$ ] or that the future spot rate will exceed the forward rate [i.e.,  $s(t+n) > f(t,n)$ ]. The forecaster does not attempt to (or is not able to) predict the magnitude of  $s(t+n) - f(t,n)$ .

The model can be formally described in terms of the probabilities of a correct forecast, conditional upon whether or not  $s(t+n) > f(t,n)$ . Let

$\gamma(t)$  be the forecaster's prediction variable where  $\gamma(t)=1$  if the forecast, made at time  $t$  is that  $s(t+n) > f(t,n)$  and  $\gamma(t)=0$  if the forecast is that  $s(t+n) \leq f(t,n)$ . The probabilities for  $\gamma(t)$  conditional upon the realized value of  $s(t+n) - f(t,n)$  are

$$p_1(t) = \text{prob} \{ \gamma(t)=0 \mid s(t+n) \leq f(t,n) \} \quad (1a)$$

and  $1-p_1(t) = \text{prob} \{ \gamma(t)=1 \mid s(t+n) \leq f(t,n) \}$

$$p_2(t) = \text{prob} \{ \gamma(t)=1 \mid s(t+n) > f(t,n) \} \quad (1b)$$

$$1-p_2(t) = \text{prob} \{ \gamma(t)=0 \mid s(t+n) > f(t,n) \}$$

Therefore  $p_1(t)$  is the conditional probability of a correct forecast, given that  $s(t+n) \leq f(t,n)$  and  $p_2(t)$  is the conditional probability of a correct forecast, given that  $s(t+n) > f(t,n)$ . Neither  $p_1(t)$  or  $p_2(t)$  depend on the level of distribution of the future spot rate. The probability of a correct forecast is assumed to be independent of the magnitude of  $s(t+n) - f(t,n)$  and only depends on whether or not  $s(t+n) > f(t,n)$ .

Merton [1981] showed that a necessary and sufficient condition for a forecaster's predictions to have no value is that  $p_1(t) + p_2(t) = 1$ . Under this condition, knowledge of the forecast will not cause an investor to change his prior estimate of the distribution of returns on the securities being evaluated. In this paper, this means the distribution of future spot rates. The existence of forecasting ability will result in  $p_1(t) + p_2(t) > 1$ . Therefore, a test of forecasting ability is to determine if  $p_1(t) + p_2(t) = 1$ .<sup>6</sup>

The nonparametric tests applied by Henriksson and Merton [1981] take advantage of the fact that the conditional probabilities of a correct forecast can be used to measure forecasting ability without requiring any assumptions regarding the distribution of future spot rates or any

particular model for security valuation. The tests examine the null hypothesis of no forecasting ability, i.e.,  $H_0: p_1(t) + p_2(t) = 1$ , where the conditional probabilities of a correct forecast,  $p_1(t)$  and  $p_2(t)$  are not known. The test determines the probability,  $P$ , that a given outcome from a sample came from a population that satisfies the null hypothesis.

Henriksson and Merton show that the null hypothesis is defined by the hypergeometric distribution:

$$P(n_1 | N_1, N_2, n) = \frac{\binom{N_1}{n_1} \binom{N_2}{n-n_1}}{\binom{N}{n}} \quad (2)$$

where  $n_1 \equiv$  number of correct forecasts, given  $s(t+n) \leq f(t,n)$ ;  $n \equiv$  number of times forecast that  $s(t+n) \leq f(t,n)$ ;  $N_1 \equiv$  number of observations where  $s(t+n) \leq f(t,n)$ ;  $N_2 \equiv$  number of observations where  $s(t+n) > f(t,n)$ ; and  $N \equiv N_1 + N_2 =$  total number of observations.

The distribution is independent of both  $p_1$  and  $p_2$ , therefore to test the null hypothesis of no forecasting ability it is not necessary to estimate either of the conditional probabilities. If the forecasts are known, all of the variables necessary for the test are directly observable. Given  $N_1$ ,  $N_2$ , and  $n$ , the distribution of  $n_1$  is determined by (2) for the null hypothesis where the feasible range for  $n_1$  is given by

$$\underline{n}_1 \equiv \max(0, n-N_2) \leq n_1 \leq \min(N_1, n) \equiv \bar{n}_1. \quad (3)$$

Equations (2) and (3) can be used to establish confidence intervals for testing the hypothesis of no forecasting ability. The appropriate criteria for evaluating forecasting ability is a one-tail test. If forecasters are rational, then it will never be true that  $p_1(t) + p_2(t) < 1$ . Small values of  $n_1$  will strictly be the result of chance, no matter how unlikely

the outcome. It seems unrealistic that a forecaster who was able to generate significant forecasting information, would not also have the ability to realize that the forecasts were systematically perverse. After all, if the forecaster's conditional probabilities of correct forecast are such that  $p_1(t) + p_2(t) < 1$ , then a strategy of doing the opposite of the forecasts will have conditional probabilities  $p_1'(t) = 1-p_1(t)$  and  $p_2'(t) = 1-p_2(t)$ . Therefore,  $p_1'(t) + p_2'(t) > 1$  and such a strategy will have value. It is just as valuable to be consistently wrong as right as long as the perversity is recognized.

In a one-tail test, the null hypothesis will be rejected with a probability confidence level of  $c$  when  $n_1 \geq x^*(c)$  where  $x^*(c)$  is determined from the solution to

$$\sum_{x \geq x^*}^{\bar{n}_1} \binom{N_1}{x} \binom{N_2}{n-x} / \binom{N}{n} = 1 - c. \quad (4)$$

It is straightforward to use the same procedure to evaluate a forecaster who either does not make a forecast in each period or who makes multiple forecasts, where the forecasts differ by the confidence of the predictions. Periods without forecasts can simply be ignored in the evaluation. When there are more than one set of forecasts, then each set can be evaluated separately, ignoring periods where the forecast does not come from the set being evaluated. To evaluate foreign exchange forecasts, it is only necessary to assume that the relationship is stationary and that for the set of forecasts being evaluated, the probability of a correct forecast only depends on whether or not  $s(t+n) > f(t,n)$ .

An example of multiple forecasts is provided by one of the foreign exchange advisory services that is evaluated. That service provides both

strong and weak forecasts of whether or not  $s(t+n) > f(t,n)$ . Therefore, it is possible to separately evaluate the strong and weak forecasts.

By focusing on the conditional frequencies of correct forecasts, it is not necessary to make any assumption about the distribution of future spot rates. Because of this,  $p_1(t)$  need not be equal to  $p_2(t)$ . This differs from the unconditional tests which require the assumption that  $p_1(t) = p_2(t)$ . For the null hypothesis of no forecasting ability, this requires that the unconditional probability of a correct forecast be equal to the probability of either of the two possible outcomes occurring (either  $s(t+n) > f(t,n)$  or  $s(t+n) \leq f(t,n)$ ) which must be assumed to be 50 percent. If one assumes that  $p_1(t) = p_2(t)$ , then the distribution of outcomes drawn from a population that satisfies the null hypothesis of no forecasting ability is the binomial distribution which can be written as

$$P(k|N,p) = \binom{N}{k} (.5)^N \quad (5)$$

where  $k$  is the number of correct predictions and  $N$  is the total number of observations.

Using (5), it is straightforward to test the joint hypothesis of no forecasting ability and that  $p_1(t) = p_2(t)$ . However, it is important to remember that such a test is a joint test and that unless  $p_1(t) = p_2(t)$ , Merton [1981] shows that an unconditional probability of a correct forecast greater than one-half,  $p(t) > .5$ , is neither a necessary nor a sufficient condition for the forecasts to have value.

One can also use (5) to test the hypothesis that the probability of  $s(t+n)$  exceeding  $f(t,n)$  is equal to 50 percent. In this case,  $k$  is the number of observations in the sample where  $s(t+n) > f(t,n)$ . This hypothesis is tested in Section IV.

#### IV. WEEKLY DATA AND SPOT-FORWARD RATE RELATIONSHIP

The techniques outlined in Section III require knowledge of the forecasts being evaluated. In this paper, the source of the forecasts are two foreign exchange advisory services. Each forecaster provided weekly advice on whether or not to hedge an exposed position with a 6 to 12 month maturity in a particular currency relative to the U.S. dollar. As the interval between forecasts is one week, this is the relevant interval for testing the hypothesis that the probability that  $s(t+n)$  will exceed  $f(t,n)$  is equal to 50 percent.

The test is run using the unconditional nonparametric test described in Section III as the null hypothesis to be tested is that each of the two possible outcomes is equally likely. The variable  $k$  in the test represents the number of actual outcomes where  $s(t+n) > f(t,n)$ . Of course, for such a test, the confidence interval will be two-tailed.

If data for forward contracts maturing each week were available, such a test would be straightforward. Unfortunately, such data is not available. Also, there is not an active secondary market for forward contracts. Therefore, it is necessary to construct a proxy for the change in the forward rate over a period of a week for a specified delivery date in the future. We want to compare  $f(t,n)$  and  $f(t+1,n-1)$  to see whether or not it was beneficial to take a hedged position in the currency at time  $t$ . Data for  $f(t,n)$  is available for contract intervals of one month, three months, six months, and one year. Data for  $f(t+1,n-1)$ , where  $n$  is measured in weeks, is constructed as follows:

$$f(t+1,n-1) = \frac{f(t+1,n)}{s(t+1)} \frac{n-1}{n} \cdot s(t+1) \quad (6)$$

The construction of  $f(t+1, n-1)$  assumes a flat term structure for the evolution of the forward rate, a potential source of error that could be quite important in parametric tests. However, in the nonparametric tests presented in this paper, we are really only interested in whether or not  $f(t+1, n-1) - f(t, n)$  is positive. The only source of error will be if the entire change is the result of events expected to take place between time  $(t+n)$  and time  $(t+n+1)$ , the period after the expiration of the forward contract under consideration. Therefore, the use of (6) to determine the sign of  $f(t+1, n-1) - f(t, n)$  will almost certainly provide an accurate estimate.

By using (6), the number of observations is quadruple the number that would be available if the shortest forward contract interval, one month, was used as the forecast period. In this paper, six month forward contracts are used for the estimation, although the results are not sensitive to the choice of contract duration.<sup>9</sup>

Using the binomial distribution described in Section III, changes in the forward rate for a specific time in the future are examined to see if the observed behavior is consistent with the hypothesis that the probability of a positive change is equal to the probability of a negative change. The test is run for nine currencies, relative to the U.S. dollar, using weekly intervals from 1977-1980. The results are shown in Table IV.1.

The null hypothesis that  $p = .5$  is rejected for the United Kingdom at the 99 percent confidence level and for Italy at the 95 percent confidence level. In addition, both Canada and Japan would reject the null hypothesis for a 90 percent confidence interval. Based on this evidence, it is clear that the results from any test that requires the assumption of unbiasedness must be suspect. Therefore, in the following tests of forecasting ability,

Table IV.1 Forward Rate Changes: Test of the Median

$$H_0: \text{prob}\{f(t+1, n-1) > f(t, n)\} = .5$$

1977 - 1980 (208 Observations)

	<u>k</u>	<u>E(p)=(k/N)</u>
Belgium	116	.56
Canada	91	.44
France	111	.53
Germany	111	.53
Italy	119	.57*
Japan	118	.57
Netherlands	113	.54
Switzerland	112	.54
United Kingdom	133	.64**

K  $\equiv$  Number of observations where  $f(t+1, n-1) > f(t, n)$

N  $\equiv$  Total number of observations

\*Reject null hypothesis with 95 percent confidence.

\*\*Reject null hypothesis with 99 percent confidence.

the focus will be on the conditional probabilities of a correct forecast, thus not requiring any assumptions concerning the distribution of either spot rates or future forward rates.

#### V TESTS OF FORECASTING ABILITY: EMPIRICAL RESULTS

The forecast of two foreign exchange advisory services,<sup>10</sup> referred to here as X and Y, are evaluated using the nonparametric procedures described in Section III for the 208 weeks of 1977-1980. Each forecaster advised weekly on whether or not to hedge an exposed position in a particular currency relative to the U.S. dollar. The forecasts, for nine different currencies, are evaluated with respect to the realized value of  $f(t+1, n-1) - f(t, n)$ , as derived in Section IV. One of the services, Y, discriminated among its forecasts by specifying hedge levels, as it had more confidence in some of its forecasts than others. Therefore, both the strong and weak forecasts are also evaluated for Service Y.

The results for the forecasts of the two services are shown in Table V.1. The table also allows for comparison of the results for the conditional tests with the unconditional tests which also do not require any assumptions regarding the distribution of spot rates or forward rates but which do implicitly assume the unbiasedness of the forward rate as predictor of future spot rates.

The results of the two different tests are quite similar for forecaster X. In both cases, only the forecasts for the Japanese yen demonstrate any predictive ability. For the conditional test, the forecasts for Japan reject the null hypothesis that  $p_1(t) + p_2(t) = 1$  with 95 percent confidence for the total period, 1977-1980, and for the second subperiod, 1979-1980. For the unconditional tests, the forecasts of the Japanese yen

Table V.1 Test of Forecasting Ability  
Two Foreign Exchange Advisory Services  
1977 1980 (208 Observations)

	Proportion Correct					
	Conditional: $E(p_1+p_2)$			Unconditional: $E(p)$		
	1977-78	1979-80	1977-80	1977-78	1979-80	1977-80
Belgium	.81	1.07	1.02	.44	.56	.50
Canada	1.03	.91	.97	.57	.47	.52
France	.98	1.06	1.02	.43	.56	.50
Germany	.95	1.10	1.05	.45	.58	.51
Italy	1.00	1.02*	1.03*	.57	.51**	.54**
Japan	1.16	1.21*	1.24*	.58	.63**	.60**
Netherlands	.83	.94	.90	.40*	.47	.44
Switzerland	1.16	.96	1.05	.59*	.47	.53
United Kingdom	.91	.96	.94	.44	.42	.43
Forecaster Y						
Belgium	.88	.81	.89	.51	.40*	.46**
Canada	1.00	1.17	1.08	.58	.59*	.59**
France	.98	.93	.92	.45**	.46	.46
Germany	1.02	.91	1.05	.63**	.45	.54*
Italy	1.10*	.99	1.14	.63**	.50	.56**
Japan	1.13*	1.02	1.13	.69*	.50	.60**
Netherlands	1.09	.92	1.03	.60*	.46	.53*
Switzerland	.94**	.95	1.10	.62**	.50	.56*
United Kingdom	1.28**	1.02	1.15	.68	.48	.58

Conditional Test  $H_0: p_1(t) + p_2(t) = 1$

Unconditional Test  $H_0: p(t) = .5$

\* Reject null hypothesis with 95 percent confidence.

\*\* Reject null hypothesis with 99 percent confidence.

reject the null hypothesis that  $p(t) = .5$  with 99 percent confidence for the same two periods.

Further evidence of the lack of forecasting ability by Service X can be found by examining the results of the two subperiods for stationarity, as shown in Table V.1. In the conditional tests, only the forecasts for Japan had estimates of  $p_1(t) + p_2(t) > 1$  for both subperiods. In the unconditional tests, only the forecasts for Japan and Italy had estimates of  $P(t) > .5$  for both subperiods, with the estimate for Italy from 1979-1980 only equal to .51.

The results from the conditional tests, however, are quite different from the results from the unconditional tests for forecaster Y. As Table V.1 shows, the forecasts for five of the nine currencies reject the null hypothesis that  $p(t) = .5$  with 95 percent confidence for the total period and six of nine reject the null hypothesis for the first subperiod, as evaluated by the unconditional test. This is in contrast to the results for the conditional tests, where none of the sets of forecasts reject the null hypothesis for the entire period, and only forecasts for two of the countries reject the null hypothesis with 95 percent confidence for the first subperiod.

An excellent example of how the assumption of unbiasedness can influence the results can be found in the evaluation of the forecasts of Service Y for Switzerland from 1977-1978. In the unconditional test, the null hypothesis that  $p(t) = .5$  is rejected with 95 percent confidence as the estimate of the unconditional probability of a correct forecast is  $E(p) = .62$ . However, in the conditional test, which does not require the assumption of unbiasedness, the estimate of  $(p_1 + p_2)$  is actually less than one,  $E(p_1 + p_2) = .94$ , clearly showing no evidence of forecasting ability.

The results for the unconditional tests for the two subperiods also demonstrate the potential problems from assuming stationarity in evaluating forecasting ability. Eight of the nine countries had estimates of  $p(t) > .5$  for 1977-1978, yet only one of the eight, Canada, had an estimate of  $p(t) > .5$  for 1979-1980. The results for the entire period, 1977-1980, are almost certainly due to the first two years. In the conditional tests, only Japan and the United Kingdom had estimates of  $(p_1+p_2) > 1$  for the two subperiods, and for both,  $E(p_1+p_2)$  was only 1.02 for the second subperiod.

In Table V.2, the strong and weak forecasts of Service Y are evaluated, using the conditional tests. It appears that the strong forecasts do outperform the weak forecasts. For the entire period, the strong forecasts had a higher estimate of  $(p_1+p_2)$  than the weak forecasts for all but one of the countries. In addition, the separation reveals some evidence of forecasting ability as the strong forecasts for the period from 1977-1980 for Italy, Japan, and the United Kingdom all reject the null hypothesis with 95 percent confidence. This is in contrast to the results for the same period for the weak forecasts as the null hypothesis could not be rejected for any of the countries. As it is possible to distinguish between the forecasts when they are made, evidence of forecasting ability in the set of strong forecasts also provides evidence of the violation of market efficiency.

One assumption required for the conditional tests is that the probability of a correct forecast not be dependent on the magnitude of  $|f(t+1,n-1) - f(t,n)|$ . This would be violated if the forecaster is able to predict periods with extreme changes better than other periods. To test for this, the sample data was split in half by the magnitude of  $|f(t+1,n-1) - f(t,n)|$ . Periods where this absolute value are greater

Table V.2

Test of Forecasting Ability  
 Different Levels of Confidence  
 Forecaster Y: 1977-1980  
 $E(p_1+p_2)$

	Forecasts	
	Strong	Weak
Belgium	.97	.81
Canada	1.07	1.10
France	1.03	.87
Germany	1.14	.92
Italy	1.15 <sup>*</sup>	1.12
Japan	1.22 <sup>**</sup>	1.03
Netherlands	1.06	1.02
Switzerland	1.10	1.08
United Kingdom	1.27 <sup>*</sup>	1.08

<sup>\*</sup> Reject null hypothesis with 95 percent confidence.

<sup>\*\*</sup> Reject null hypothesis with 99 percent confidence.

than the sample median are separated from those below the median. The results of this test are shown in Table V.3.

There does not appear to be much of a difference in the results for large magnitude changes and small magnitude changes for Service X. Four countries have higher estimates of  $(p_1+p_2)$  for the total period in the small magnitude sample and five countries have higher estimates in the large magnitude sample. For the first subperiod, the estimate was higher for the large magnitude sample for only three countries, but was higher for seven countries in the second subperiod. France rejected the null hypothesis with 95 percent confidence in the large magnitude sample yet had an estimate below one,  $E(p_1+p_2) = .85$ , in the small magnitude sample. On the other hand, the forecasts for Japan could not reject the null hypothesis in the large magnitude sample and yet could reject it with 99 percent confidence in the small magnitude sample.

In contrast to the forecasts of Service X, the forecasts of Service Y do appear to perform better in the large magnitude sample than in the small magnitude sample. Seven of the nine countries had higher estimates of  $(p_1+p_2)$  for the total period, 1977-1980, in the large magnitude sample than the small magnitude sample. In the first subperiod, seven of the countries had higher estimates for the large magnitude sample and six of the countries had higher estimates for the large magnitude sample in the second subperiod.

In addition, the forecasts for Italy, Switzerland, and the United Kingdom in the large magnitude sample all reject the null hypothesis of no forecasting ability with 95 percent confidence while none of the sets of forecasts reject the null hypothesis in the small magnitude sample. If a forecaster is more likely to be able to predict the large magnitude changes

Table V.3

Test of Forecasting Ability  
 Sample Split by Magnitude of  
 $f(t+1, n-1) - f(t, n)$

$$E(p_1 + p_2)$$

Forecaster X	Small Magnitudes			Large Magnitudes		
	1977-78	1979-80	1977-80	1977-78	1979-80	1977-80
Belgium	.99	.93	1.02	.52	1.20*	1.06
Canada	1.06	.87	.99	.90	.93	.94
France	.85	.91	.85	1.17	1.18	1.20*
Germany	1.06	1.04	1.07	.77	1.14	1.03
Italy	.93	1.02	.98	1.06	1.04	1.10
Japan	1.27*	1.24	1.33**	1.01	1.19	1.16
Netherlands	.95	.78	.88	.61	1.04	.91
Switzerland	1.04	.92	1.01	1.38*	1.02	1.10
United Kingdom	1.02	.96	1.00	.76	.97	.89
Forecaster Y						
Belgium	.83	.87	.88	.97	.77	.90
Canada	1.00	1.02	1.01	1.00	1.19	1.13
France	.83	.89	.83	1.16	.93	.99
Germany	.97	.75	.98	1.10	1.01	1.12
Italy	1.06	.98	1.11	1.22	.99	1.19*
Japan	1.11	1.01	1.13	1.15	1.01	1.12
Netherlands	1.07	1.03	1.07	1.14	.86	.99
Switzerland	.98	.79	.97	.89	1.06	1.19*
United Kingdom	1.09	1.06	1.09	1.55**	.99	1.22*

\* Reject null hypothesis with 95 percent confidence.

\*\* Reject null hypothesis with 99 percent confidence.

than the smaller changes, then the results of the tests using all outcomes will be biased against finding forecasting ability. It is certainly true that it is more valuable to be able to predict large changes than small changes. A strategy that follows the predictions of a forecaster who has forecasting ability for periods with large magnitude outcomes, but not for periods with small magnitude outcomes, will have value because the impact of the periods with small changes will be minimal in comparison with the impact of the periods with large changes as the costs from errors in periods with small changes will be small. Thus, the results for the forecasters in the large magnitude sample for Service Y, as shown in Table V.3, provide additional evidence of the violation of market efficiency.

In addition, Table V.4 shows that the evidence of forecasting ability found in the strong forecasters of Service Y for Italy, Japan, and the United Kingdom was the result of forecasters for periods with large magnitude changes. For those three countries, the forecasts that Service Y had most confidence in show evidence of successfully forecasting the periods with the largest changes in the forward rate.

## VI. CONCLUSIONS

Using the nonparametric techniques developed by Henriksson and Merton [1981], the hypothesis of forward foreign exchange market efficiency has been tested through the evaluation of the forecasting ability of two foreign exchange advisory services. Unlike previous tests, this methodology does not require any assumptions regarding the relationship between the forward rate and the corresponding expected future spot rate. Because the true model of forward rate determination is not known, none of the previous tests

Table V.4                      Test of Forecasting Ability  
 Different Levels of Confidence  
 Sample Split by Magnitude  
 1977 - 1980

	Small Magnitudes		Large Magnitudes	
	Forecast		Forecast	
	<u>Strong</u>	<u>Weak</u>	<u>Strong</u>	<u>Weak</u>
Belgium	.94	.82	1.00	.81
Canada	1.03	.94	1.09	1.31*
France	.99	.79	1.09	.90
Germany	1.05	.83	1.24	1.00
Italy	1.09	1.11	1.24*	1.17
Japan	1.19	1.04	1.27*	1.03
Netherlands	1.14	1.04	.93	.99
Switzerland	.95	1.01	1.24	1.10
United Kingdom	1.17	1.03	1.36*	1.12

\* Reject null hypothesis with 95 percent confidence.

\*\* Reject null hypothesis with 99 percent confidence.

can be considered a definitive examination of forward market efficiency. The methodology of this paper provides the first real opportunity to examine the information efficiency of the forward market.

While the technique does not require information regarding the magnitude of gains from forecasting, it thus does not provide direct measure of its value. However, Merton [1981] has shown that successful timing advice can be valued as a put option -- in this case of the return on a foreign treasury bill with a striking price of the return on a U.S. treasury bill.

In addition, because of the nonparametric nature of the tests, it is possible to evaluate weekly forecasts, substantially increasing the number of available independent observations. The empirical results show some evidence of forecasting ability on the part of one of the services, evidence that violates the hypothesis of market efficiency.

Using the entire sample for the period from 1977-1980, it was not possible to detect significant forecasting ability. However, when the forecasts of which one of the services was most confident were used, significant forecasting ability for the lira, yen, and pound sterling relative to the U.S. dollar was shown. That service also appeared to be better at forecasting large changes in the forward exchange rate than smaller changes. In fact, the successful forecasting ability, reflected in the predictions of which the service was most confident, was the result of predictions for periods with the largest changes. The successful forecasting ability of the strongest forecasts seems to correspond to successfully predicting the largest changes.

Most previous tests of forecasting ability have required the assumption that the forward rate is an unbiased estimate of the expected future spot

rate. In this paper, evidence rejecting this assumption was presented and it was shown how this assumption can change the empirical results.

## FOOTNOTES

1. Outside assets are defined as nominal assets that are issued by governments and are not viewed by the residents of the country as a liability.
2. See for example, Giddy and Dufey [1975], Levich [1978], Westerfield [1977], and McFarland, Pettit, and Sung [1982].
3. See for example, Levich [1981, 1982a, 1982b].
4. The validity of the unconditional probability of a correct forecast as a measure of forecasting ability has been the subject of much debate in recent issues of Euromoney. It is valid only if the assumptions listed above are valid.
5. In addition, unless the sample size is quite large, there will be periods where there will be many more of one of the outcomes than the other, even if the ex ante probabilities of each of the two outcomes is equal. If forecasting ability is evaluated using the unconditional test for such a period, a forecaster who always makes the same prediction may appear to have forecasting ability and yet it is obvious that such forecasts have no value.
6. An analogy to the type of forecast modelled, suggested by Arnold Barnett, is the problem of a forecaster faced with a bin full of apples and oranges trying to predict which type of fruit will be drawn next. If the forecaster has no information except the number of each type of fruit in the bin, then the probability of selecting either an apple or an orange will be independent of the forecaster's prediction and  $p_1(t) + p_2(t) = 1$ . where the probabilities are conditioned on whether an apple or an orange was selected.

7. Because of hypergeometric distribution is discrete the strict equality of (4), will usually not be obtainable. Therefore, in (4),  $x^*$  should be interpreted as the lowest value of  $x$  for which the summation does not exceed  $(1-c)$ .
8. The maturity of the exposure that is being evaluated is assumed by the forecaster to be approximately six months. However, as forecasts can be updated weekly, the forecasts should focus on developments over the next week that will effect this exposure. The decision can be thought of as choosing between a U.S. treasury bill with six months to maturity when the investment horizon is one week.
9. The tests presented in this paper were also run for a few of the currencies using three-month and one-year forward contracts with no qualitative difference in the results.
10. The forecasts for the two services were obtained from a corporation with large foreign exchange exposure in their accounts receivable. The company subscribed to the two services and provided the information with the stipulation that the name of the company and the two services would not be revealed. As previously mentioned, the forecasts were provided weekly and focused on the relationship between the forward rate and the actual future spot rate.

#### REFERENCES

1. Bilson, J., 1981. "The 'Speculative Efficiency' Hypothesis," Journal of Business 54 (July): 435-451.
2. Breeden, D.T., 1979. "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," Journal of Financial Economics 7 (September): 265-296.
3. Dornbusch, R., 1979. "Monetary Policy Under Exchange Rate Flexibility," in Managed Exchange Rate Flexibility, Federal Reserve Bank of Boston, Conference Vol. No. 20.
4. Fama, E.F. and A. Farber, 1979. "Money, Bonds, and Foreign Exchange," American Economic Review 69 (September): 639-649.
5. Frankel, J.A., 1979. "The Diversifiability of Exchange Risk," Journal of International Business Studies 9: 379-393.
6. Giddy, I.H. and G. Dufey, 1975. "The Random Behavior of Flexible Exchange Rates," Journal of International Business Studies 6 (Spring): 1-32.
7. Grauer, F.L.A., R.H. Litzenger and R.E. Stehle, 1976. "Sharing Rules and Equilibrium in an International Capital Market Under Uncertainty," Journal of Financial Economics 3 (June): 233-256.
8. Hansen, L.P. and R.J. Hodrick, 1980. "Forward Exchange Rates as Optimal Redictors of Future Spot Rates: An Econometric Analysis," Journal of Political Economy 88 (October): 829-853.
9. Henriksson, R.D. and R.C. Merton, 1981. "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills," Journal of Business 54 (October): 513-533.
10. Isard, P., 1978. "Exchange Rate Determination: A Survey of Popular Views and Recent Models," Princeton Studies in International Finance No. 42.
11. Kohlhagen, S.W., 1978. The Behavior of Foreign Exchange Markets - A Critical Survey of the Empirical Literature, New York University Monograph Series in Finance and Economics, No. 1978-3.
12. Kouri, P.J.K., 1977. "International Investment and Interest Rate Linkages Under Flexible Exchange Rates," in R.Z. Aliber, ed., The Political Economy of Monetary Reform.
13. Levich, R.M., 1978. "Tests of Forecasting Models and Market Efficiency in the International Money Market," in J.A. Frenkel and H.G. Johnson, eds., The Economics of Exchange Rates.

14. Levich, R.M., 1979. "The Efficiency of Markets for Foreign Exchange: A Review and Extension," in D.R. Lessard, ed., International Financial Management: Theory and Application
15. Levich, R.M., 1981a. "How to Compare Chance With Forecasting Expertise," Euromoney (August): 61-68.
16. Levich, R.M., 1982a. "Evaluating the Performance of Professional Forecasters," in Richard Ensor, ed., Management of Foreign Exchange Risk, 2nd Edition. London: Euromoney Publications Ltd.
17. Levich, R.M., 1982b. "Evaluating the Performance of Professional Foreign Exchange Forecasters: An Update Report," Euromoney (August).
18. McFarland, J.W., R.R. Pettit, and S.K. Sung, 1982. "The Distribution of Foreign Exchange Price Changes: Trading Day Effects and Risk Measurement," Journal of Finance 37 (June): 693-715.
19. Merton, R.C., 1973. "An Intertemporal Capital Asset Pricing Model," Econometrica 41: 867-887.
20. Merton, R.C., 1980. "On Estimating the Expected Return on the Market: An Exploratory Investigation," Journal of Financial Economics 8 (December): 323-361.
21. Merton, R.C., 1981. "On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts," Journal of Business 54 (July): 363-406.
22. Stulz, R.M., 1982. "The Forward Exchange Rate and Macroeconomics," Journal of International Economics, (May).
23. Westerfield, J.M., 1977. "An Examination of Foreign Exchange Risk Under Fixed and Floating Rate Regimes," Journal of International Economics 7 (May): 181-200.