# Forecasting Emerging Market Returns Using Neural Networks

CAMPBELL R. HARVEY, KIRSTEN E. TRAVERS, AND MICHAEL J. COSTA*

CAMPBELL R. HARVEY is the J. Paul Stricht Professor of international business at the Fuqua School of Business at Duke University, and is an editor of *Emerging Markets Quarterly*.

KIRSTEN E. TRAVERS is an associate at Watch Hill Capital in Atlanta.

MICHAEL J. COSTA is with Pine Grove Associates in Summit, New Jersey.

Given the extreme volatility of emerging market returns, it would seem difficult to uncover any predictability. Despite this volatility, Harvey [1995a] documents more predictability in emerging market returns than developed market returns. One hypothesis is that emerging equity markets are less informationally efficient than developed markets and the degree of inefficiency can be exploited in a forecasting model. Bekaert and Harvey [1995, 1997] argue that the predictability is complicated in that it is unstable through time.

Emerging markets provide a good testing ground for the viability of nonlinear forecasting techniques. Indeed, there has been much interest in the last decade in the application of neural networks to finance, in particular stock price prediction and selection. In this article, we investigate if returns in emerging market returns can be better forecast using neural networks rather than linear prediction models.

Linear systems are very simple, and it is a great simplification to assume that a process as complicated as financial markets could be driven by a system whose relationships are so easy to express. The advantage of linear models is their simplicity and their ease of use. However, a model is only useful as long as its predictions do not deviate too far from the outcome of the underlying process. It is this trade-off between simplicity of a model and accuracy of predictions which is of great importance to the researcher trying to predict market returns in real time.

Linear regression forecasting models have demonstrated their usefulness in predicting returns in both developed markets (see Harvey [1991]) and emerging markets (see Harvey [1995]). A good linear regression model can correctly predict direction in the market over 55%-65% of the time. However, it is reasonable to assert that many of the factors we believe drive the financial markets may not be related by linear functions. This leads us to ask what methods are available to build a non-linear model and see if it works better; in the sense that it may more accurately forecast the returns of a stock market index over a period of several years.

Nonlinear models are much more difficult to devise, partly due to the fact that there are many more nonlinear functions than linear ones, and they are thus more difficult to specify. Many nonlinear functions may have graphs that look similarly curved, whereas linear functions are straight lines, easy to identify by the unique axis intercept and constant slope.

Recent research, however, has established some methods of identifying nonlinear models such as nonlinear regression, parametric models such as Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and nonlinear volatility models and nonparametric models. One such nonparametric method is the use of neural networks. These are systems, first devised in research into artificial intelligence, in which a computer learns the nonlinear relationship between independent and dependent variables by being given large quantities of training data It is then hoped that when given out-of-sample data they can predict the outcome of a nonlinear function more accurately than linear regression. They use a non-parametric method of forecasting; that is, the underlying non-linear function is neither prescribed nor predicted explicitly. Thus, the model is not limited to a restrictive list of non-linear functions. This can be contrasted with linear regression, where an initial presumption is that the underlying relationship is of the form y = a + bx + error (where a is an intercept and b is the slope coefficient).

The plan for our article is the following. First, we explain how such a neural network is constructed. Then we investigate the accuracy of such a non-linear model by comparing the performance of a strategy for investing in nine emerging markets implied by a neural net to both that of the more traditional trading strategy of buy-and-hold and an investment strategy implied by linear regression analysis in the 1992 to June 1997 period. Finally, we apply the model to a hold-out sample: July 1997-December 1998 for one country, Korea. This country experienced extreme volatility during the Asian crisis and this sample provides a very challenging test for any prediction model.

## NEURAL NETS: THE SETTING

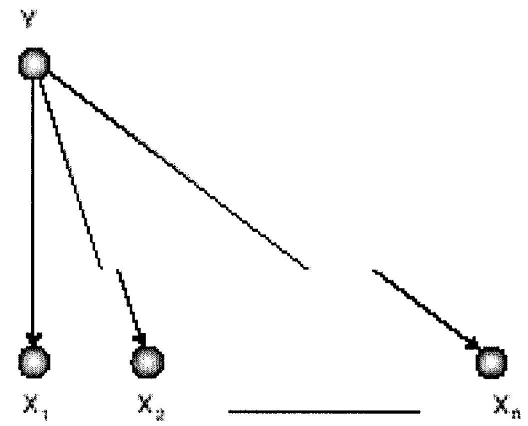
### Linear Regression Benchmark

Recall that linear regression finds the best straight-line fit between input (independent) variables and an output (dependent) variable. That is, given a time-series of inputs $X_i$, i = 1,...,n, it establishes weights $w_i$ such that

$$Y = \Sigma_i w_i X_i + \text{error}$$

where the weights are constants so that the error is minimized. Usually, we minimize the squared errors though it is possible to minimize other functions of the error like absolute value. This construction finds the line of best fit through the points $X_i$. The structure of n inputs and one output is shown in Exhibit 1.

**EXHIBIT 1**
**The Line of Best Fit**




### How Neural Networks Work

We will outline a method that trains the neural net to learn a non-linear relationship between inputs and output, by applying non-linear smoothing functions to the linear combinations of inputs given above. The method is best described using the following example of an *AND-gate.*

#### *Example: An AND-Gate*

Consider the simple nonlinear map taking two inputs, each taking the value 0 or 1, to one output that has value 0 unless both the first AND the second inputs are equal to 1. This can be thought of as a binary switch Y that is either ON or OFF.

This nonlinear function is given by the product of the two inputs,

$$Y = X_1 * X_2$$

We shall try to model this function by linear regression and by using neural networks. The purpose is to find the best approximation, which could then be used to forecast the results given two random inputs of 0 or 1.

The linear regression model can be calculated by hand, or by using a regression software package. The best linear approximation is given by:

$$Y = 0.5 X_1 + 0.5 X_2 - 0.25$$

with a root mean squared error of 0.5. The predictions of the linear model are detailed in Exhibit 2.

## EXHIBIT 2
**Performance of Neural Net Approximation**

| $X_1$ | $X_2$ | Y | Neural Net Approximation $Y_t$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.25 |
| 0 | 1 | 0 | 0.25 |
| 1 | 0 | 0 | 0.25 |
| 1 | 1 | 1 | 0.75 |

### Construction of the Neural Net

Let us see how the neural network can learn the pattern and more closely predict the in-sample results. Exhibit 3 illustrates how n inputs are combined in the neural network to give another layer of what we shall call nodes, which are then also combined to give the output Y.

In our example, we have just two inputs $X_1$, $X_2$, and one output, Y. In between, we place one extra hidden layer containing two nodes, $H_1$ and $H_2$. Nonlinear functions will relate the different layers, as in Exhibit 4.

As in the linear regression, we assign each input $X_i$ a weight $w_i$. Then the problem is as for the linear regression above: to find weights for the nodes such that the inputs may be weighted and combined (but this time nonlinearly) to best approximate the output. We do this in two stages, since there are two layers of nodes.

The inputs are binary: either 0 (off) or 1 (on). Then the hidden layer has two nodes, which are set to be:

$$H_1 = g(w_{11} X_1 + w_{12} X_2)$$
$$H_2 = g(w_{21} X_1 + w_{22} X_2)$$

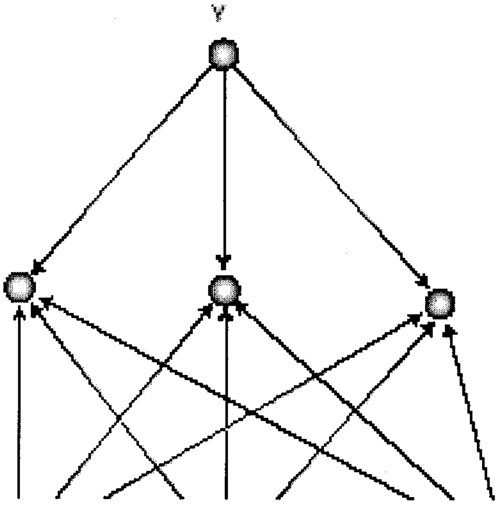
where g is a smoothing non-linear function shown in Exhibit 5.

For each time t, the output is a linear combination of the hidden layer variables $H_1$ and $H_2$, with then yet another non-linear smoothing function h applied to that:

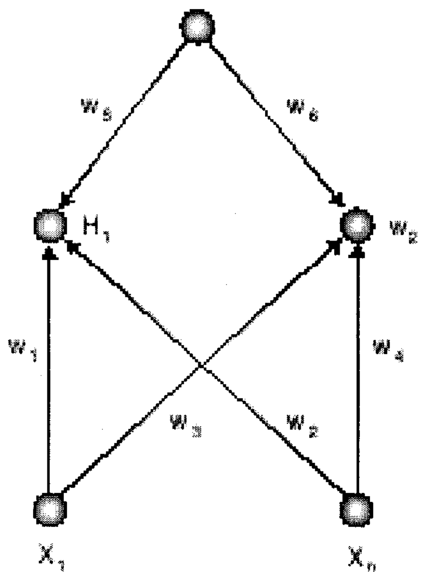
$$Y_t = h(a_1 H1 + a_2 H_2)$$

So the inputs $X_1$, $X_2$ and the output Y are fixed, and we would like the program to learn from data that we give it what the correct weights are such that Y has the correct value corresponding to the inputs. The fact that such weights exist is ensured by the *Universal Approximation Property* (White [1992]). The property says that any continuous non-linear function $Y(X_1,..,X_n)$ can be approximated as above, with one hidden layer, to an arbitrary degree of accuracy with a suitable number of nodes $H_i$ in the hidden layer.

## EXHIBIT 3
**Building a Neural Network**

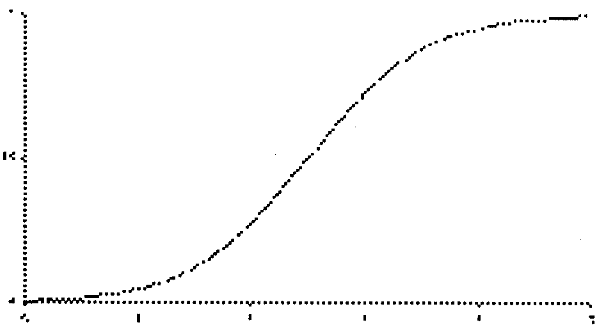



## EXHIBIT 4
**Hidden Layers**




This property guarantees that our method will yield an approximate solution that is as close as we like to the underlying function but does not tell us how to construct such an approximation. So the question remains: how does the network learn the underlying nonlinear relationship; how do we find such a $Y_t$?

## EXHIBIT 5
**The Smoothing Non-Linear Function**



### Training of the Neural Net

Training uses a method called **error-back propagation.** The neural net starts by assigning arbitrary weights to the input variables. Then it runs through all the input data points (called "training data") once, and gives a list of outputs. Then a comparison is made between the desired outputs $Y_t$ and the net's output for each time t. Then the mean-squared error

$$\Sigma_t \, |Y_t - h(\Sigma_k \, a_k \, g(\underline{w}^T.\underline{X}))|^2$$

is minimized by using the gradient method with a certain prescribed learning rate. This is done as follows. The n-dimensional gradient is computed and the weights adjusted to incrementally change the output in the direction of steepest descent, that is, in the direction in which the mean-squared error's derivative is negative with greatest absolute value. This means that the error decreases most rapidly in this direction, just as the quickest way to descend a smooth mountain to the valley is to always go down in the direction in which the slope is the steepest. The learning rate determines the rate of the error's convergence to zero.

### Testing of the Neural Net

Once the neural net has trained on a data set and can predict the output to a required degree of accuracy, it is then tested on new out-of-sample data. It is hoped that the trained network can predict the outcome $Y_t$ by substituting new unseen inputs $X_i$ into the learned approximation function.

The resulting approximation function $Y_t$ that our neural network program gave for the AND-gate example is shown in Exhibit 6. It is clear that the neural network does a far better job of approximating the nonlinear function given by the AND-gate than the linear regression method, with a much lower root mean squared error. It is hoped that this in-sample approximation would help out-of-sample predictions of the underlying nonlinear relationship.

## EXHIBIT 6
**Performance of Neural Net Approximation**

| $X_1$ | $X_2$ | Y | Neural Net Approximation $Y_t$ |
|---|---|---|---|
| 0 | 0 | 0 | 0.0009 |
| 0 | 1 | 0 | 0.0319 |
| 1 | 0 | 0 | 0.0290 |
| 1 | 1 | 1 | 0.9303 |

## RESULTS

### Data

The data was extracted from the International Finance Corporation's (IFC) Emerging Markets Data Base (EMDB). This database contains weekly data for thirty-three emerging market countries and four composites. Data included total return and three ratios: price-to-earnings (PE), price-to-book value (PBV), and dividend yield (YLD). Returns for the U.S. for the same period were also used. While the data contained data from 1989, data was not available for many countries until a much later date. To permit comparison between countries, only countries whose data sets permitted analysis for 1992-1997, inclusive were included. A total of nine countries' data series were analyzed. We use returns converted to U.S. dollar terms.

The analysis procedure was intended to provide a comparison of a neural network model to both a passive strategy (buy and hold) and an active strategy determined by the results of regression analysis.

### Regression Procedure

The allowable data for the regression analysis includes PE, PBV, YLD, four lags of country return, and one lag of U.S. return. The data is analyzed as an expanding window — all data prior to the in-sample forecast range is used; i.e., if attempting to forecast returns for 1996, we use data through 1994, test in-sample for 1995, and test out-of-sample for 1996. That is, the regression coefficients are updated every year.

Each regression begins using all available variables. The variables are then removed sequentially, removing the variable with the lowest F-statistic each time; until all remaining variables are significant at least at the 90% confidence level. The resulting formula is then used for forecasting both in- and out-of-sample. While this procedure is a classic data snooping method, it establishing a meaningful benchmark for us.

### Neural Network Procedure

The allowable data for the neural network analysis is restricted to price data: four lags of country data and one lag of U.S. return. The data is analyzed as an expanding window, in the same manner as for the regression procedure.

The method of determining the best neural network analysis differes from the method for regression analysis because methods of significance testing are only now being developed for neural network models. The objective is to generate a neural network that has the greatest modified direction (see later) for the in-sample data. This model is then used for the out-of-sample period without bias; that is, regardless of the results of out-of-sample testing, the model with the best in-sample performance is used. Thus we attempt to strike a compromise between the potential for overfitting data and the absence of significance testing.

### Analysis of Forecasts

The regression equations and neural network models are used to forecast country returns out-of-sample. The decision rule used for asset allocation is to invest in the country index if the forecast for a country return was greater than that for the risk-free U.S. Treasury bill T-bill and to invest in the T-bill otherwise.

Three performance measures are used for the two active strategies:

- *Total Return ($)*: The return per dollar invested in the strategy during the time period. It might be argued that the total return is the most important statistic, since it measures the profitability of following a particular investment strategy. As is often the case, however, in market return series, correctly forecasting a few periods (or even one period) can overwhelm the impact of all other forecasts in this measure.
- *Direction (% between 0 and 100):* The number of times a model correctly forecasts market direction divided by total forecasts. This measure eliminates the magnitude of returns on performance measurement, and thus removes some of the randomness present in the Total Return measure. However it suffers from the stopped watch problem. That is, if a strategy forecasts an up market most of the time and if the market happen to have been up most of the time, then it is likely that the strategy will correctly forecast market direction more effectively than a random guess (50%).
- *Modified Direction, Merton Measure (% between -100 and 100):* The number of times the model correctly forecasts a market return greater than T-bills plus the number of time the market correctly forecasts a market return greater than T-bills, weighted by the respective number of up and down forecasts, plus one (see Merton, [1981]) This measure eliminates both the magnitude problem and the stopped watch problem. It can be seen that a strategy that always forecasts a positive excess return will receive a 100% score for the positive actual excess return and a 0% score for the negative realized excess returns which, after subtracting one, results in a score of 0%. Of course, a strategy that is particularly bad could have a negative Modified Direction score.

### Results

Results for each country are summarized in Exhibit 7. The net strategy outperforms the buy and hold strategy in forty-four of the fifty-four country years (that is, nine countries and six years for each country). The net strategy outperforms the regression in forty of the fifty-four country years. Exhibit 7 also reports a summary of the annual volatility. The net strategy produces a lower volatility than the buy and hold in fifty-four of fifty-four country years. This means that in forty-four of the fifty-four country years the net produces higher returns and lower volatility. In the other ten country years the returns are lower but the volatility is also lower. We also report the volatility of the net versus the regression based model. The exhibit suggests that the regression model often produces lower volatility than the net. However, the reason for this is simple. The regression forecasts are so imprecise that often the algorithm will recommend holding cash the entire year (which has low volatility). The cumulative buy and hold returns are presented in Exhibits 8-16.

## EXHIBIT 7
**U.S. Dollar Returns to Various Strategies**

| | 1992 Regression | 1992 Neural Net | 1992 Passive | 1993 Regression | 1993 Neural Net | 1993 Passive |
|---|---|---|---|---|---|---|
| Argentina | -21.1% | -19.1% | -27.2% | 72.4% | 60.2% | 72.1% |
| Brazil | -11.2% | 17.1% | -19.5% | 22.3% | 126.7% | 112.5% |
| Chile | 31.0% | 48.4% | 41.6% | 61.0% | 60.9% | 41.6% |
| Columbia | 39.0% | 10.5% | 5.0% | 59.7% | 31.9% | 40.0% |
| Korea | 26.2% | 139.9% | 83.0% | 24.8% | 36.2% | 24.7% |
| Malaysia | 36.9% | 47.6% | 45.7% | 26.8% | 74.1% | 92.5% |
| Mexico | 20.5% | 17.1% | 14.4% | 51.7% | 55.6% | 49.4% |
| Taiwan | -21.1% | -19.8% | -26.5% | 72.4% | 70.9% | 94.3% |
| Thailand | 62.6% | 63.1% | 40.9% | 53.6% | 46.3% | 90.7% |

| | 1994 Regression | 1994 Neural Net | 1994 Passive | 1995 Regression | 1995 Neural Net | 1995 Passive |
|---|---|---|---|---|---|---|
| Argentina | -14.6% | 36.9% | -24.3% | 5.9% | 47.9% | 19.1% |
| Brazil | -5.3% | 22.0% | 29.7% | -25.6% | -0.2% | -10.4% |
| Chile | 37.6% | 19.1% | 32.4% | 9.2% | -4.9% | 0.6% |
| Columbia | 58.4% | 71.0% | 32.0% | 14.4% | -0.7% | -28.7% |
| Korea | 22.2% | 27.8% | 7.7% | 27.8% | 19.5% | -2.8% |
| Malaysia | -17.3% | -9.3% | -19.1% | 5.9% | 25.8% | 19.2% |
| Mexico | 23.2% | -31.9% | -47.7% | 38.7% | 20.0% | -6.9% |
| Taiwan | -14.6% | 18.0% | 15.3% | 5.9% | -6.8% | -30.2% |
| Thailand | 26.3% | 21.1% | -4.8% | 9.0% | 19.3% | 15.7% |

| | 1996 Regression | 1996 Neural Net | 1996 Passive | 1997 Regression | 1997 Neural Net | 1997 Passive |
|---|---|---|---|---|---|---|
| Argentina | 19.3% | 13.7% | 15.5% | 17.2% | 22.3% | 14.7% |
| Brazil | 27.8% | 35.1% | 24.0% | 26.3% | 42.1% | 21.9% |
| Chile | -3.6% | 6.4% | -13.2% | 19.0% | 13.8% | 0.4% |
| Columbia | 14.5% | 22.1% | 10.2% | 32.9% | 21.5% | 17.2% |
| Korea | 27.3% | 10.4% | -36.6% | 28.6% | 17.2% | -69.1% |
| Malaysia | 15.9% | 19.5% | 16.8% | -72.5% | -67.7% | -72.9% |
| Mexico | 29.0% | 7.4% | 4.8% | 44.7% | 22.3% | 14.7% |
| Taiwan | 19.3% | 22.9% | 43.1% | 17.2% | 3.9% | -7.7% |
| Thailand | -3.5% | -5.2% | -44.7% | -67.6% | -39.3% | -78.0% |

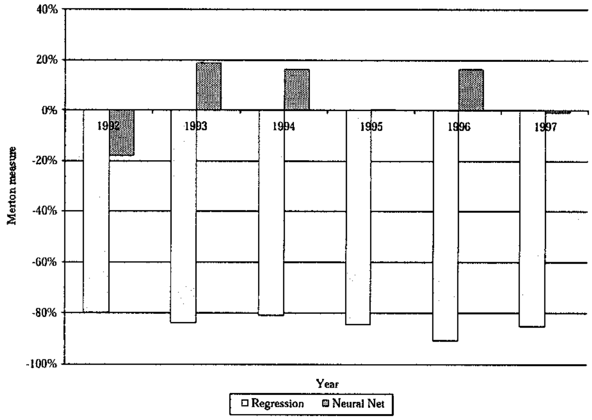
| | Net Beats Passive (wins/years) | Net Beats Regression (wins/years) | Net Lower Vol than Passive (wins/years) | Net Lower Vol than Regression (wins/years) |
|---|---|---|---|---|
| Argentina | 4/6 | 6/6 | 6/6 | 5/6 |
| Brazil | 5/6 | 6/6 | 6/6 | 1/6 |
| Chile | 4/6 | 5/6 | 6/6 | 2/6 |
| Columbia | 5/6 | 2/6 | 6/6 | 6/6 |
| Korea | 6/6 | 6/6 | 6/6 | 1/6 |
| Malaysia | 5/6 | 6/6 | 6/6 | 4/6 |
| Mexico | 6/6 | 1/6 | 6/6 | 1/6 |
| Taiwan | 4/6 | 4/6 | 6/6 | 0/6 |
| Thailand | 5/6 | 4/6 | 6/6 | 1/6 |
| Neural Net Wins/Total | 44/54 | 40/54 | 54/54 | 21/54 |
| Neural Net Win Rate | 81% | 74% | 74% | 39% |

The Modified Direction analysis is contained in Exhibits 17-25. The neutral networks outperform both the buy and hold and the regression benchmark. For example, the average Modified Direction measure (across all nine countries) is positive for all years. Looking across time, the Modified Direction measure is positive on average for all countries except for Brazil and Taiwan. The linear regression performs poorly.

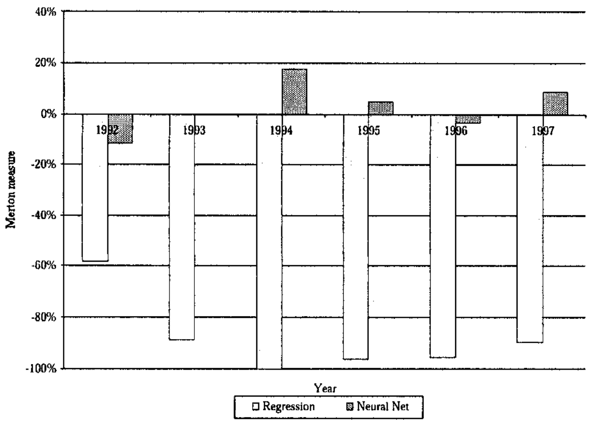
### Performance During the Asian Financial Crisis: June 1997-March 1999

Since performing our complete analysis of the nine emerging markets through 1997, we thought it would be of interest to study in more detail a country

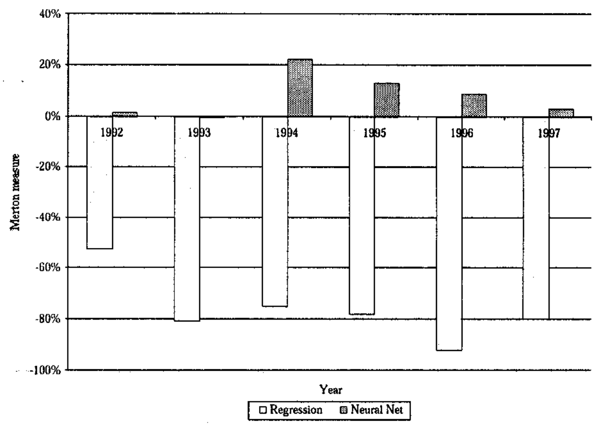
## EXHIBIT 10
**Cumulative Return on $1 Invested in Chile**



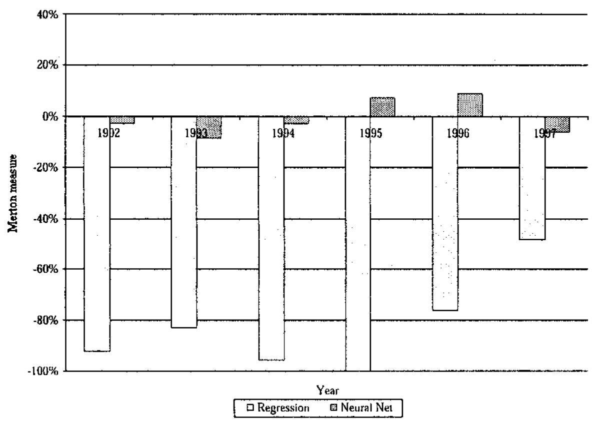

## EXHIBIT 8
**Cumulative Return on $1 Invested in Argentina**



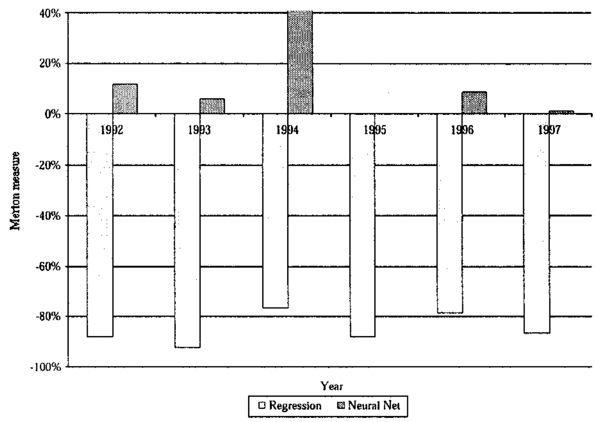

## EXHIBIT 11
**Cumulative Return on $1 Invested in Colombia**



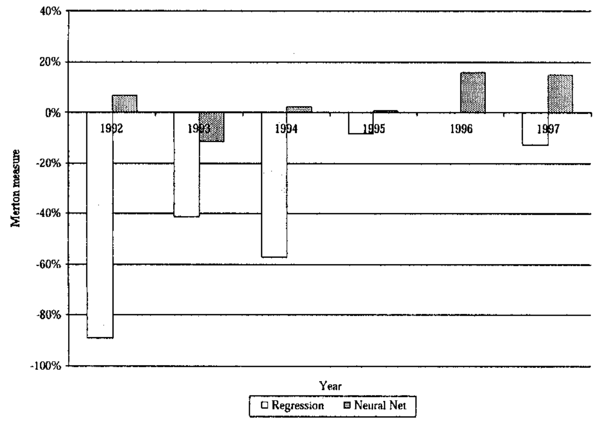

## EXHIBIT 9
**Cumulative Return on $1 Invested in Brazil**



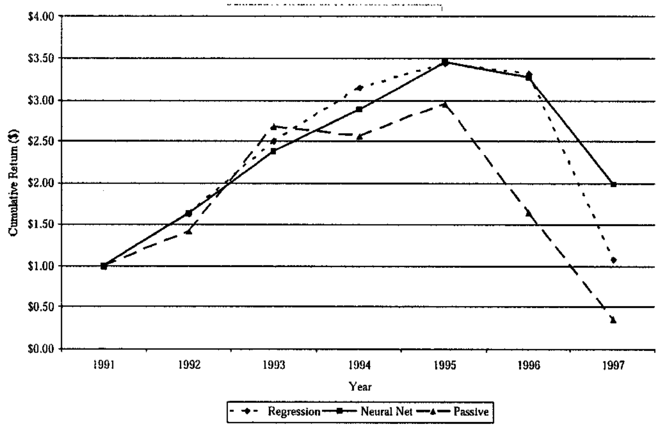

## EXHIBIT 12
**Cumulative Return on $1 Invested in Korea**



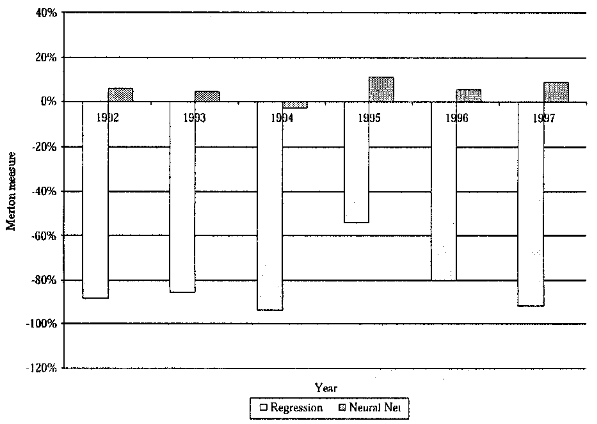
# EXHIBIT 13
## Cumulative Return on $1 Invested in Malaysia



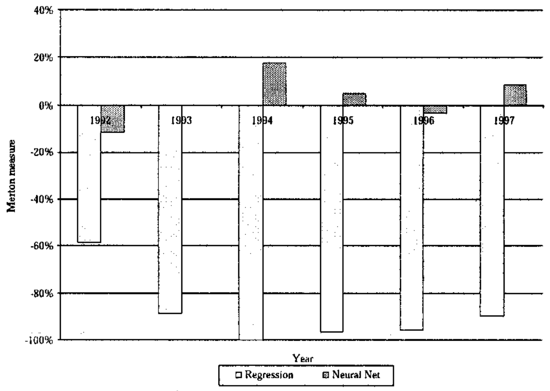

# EXHIBIT 16
## Cumulative Return on $1 Invested in Thailand



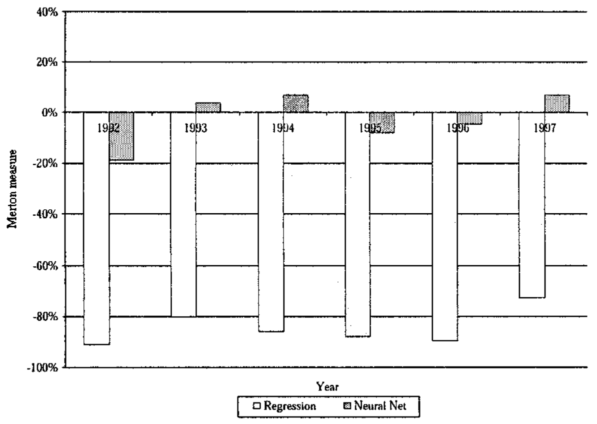

# EXHIBIT 14
## Cumulative Return on $1 Invested in Mexico



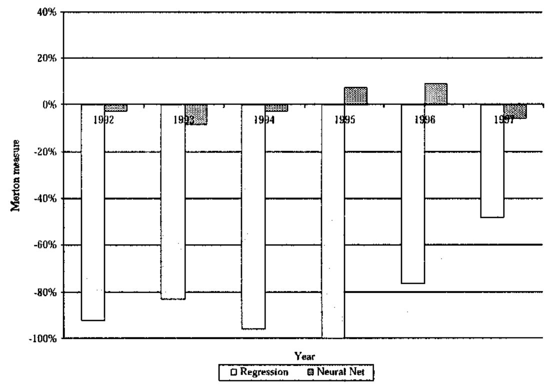

# EXHIBIT 17
## Modified Direction Measure for Argentina



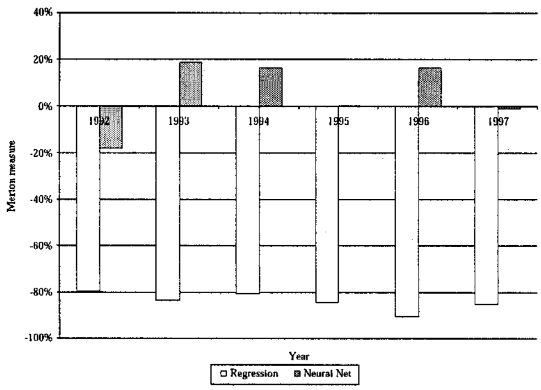

# EXHIBIT 15
## Cumulative Return on $1 Invested in Taiwan



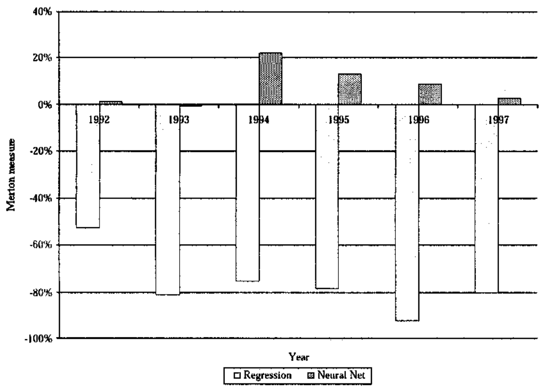

# EXHIBIT 18
## Modified Direction Measure for Brazil



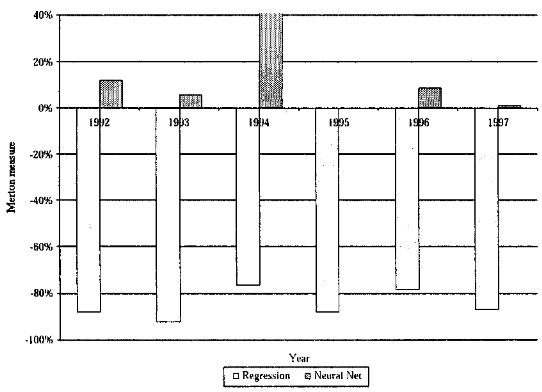
# EXHIBIT 19
## Modified Direction Measure for Chile



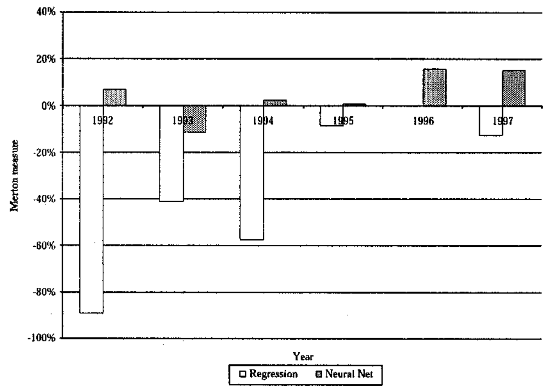

# EXHIBIT 20
## Modified Direction Measure for Colombia



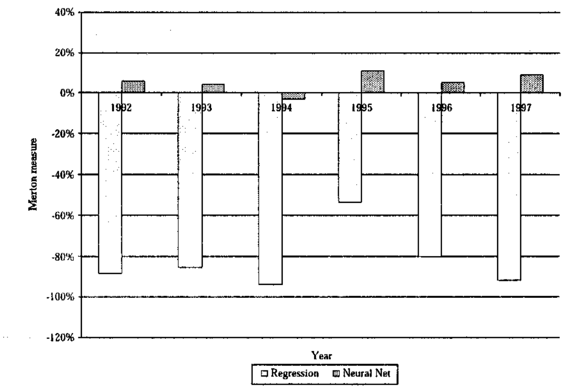

# EXHIBIT 21
## Modified Direction Measure for Korea



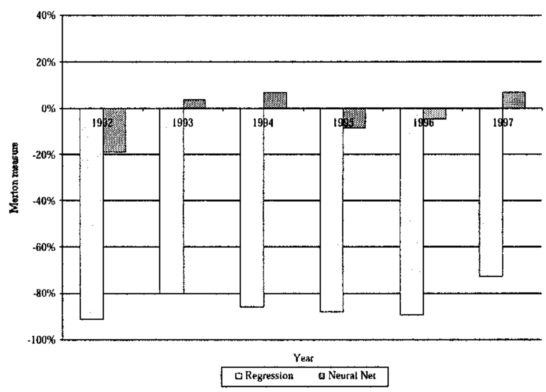

# EXHIBIT 22
## Modified Direction Measure for Malaysia



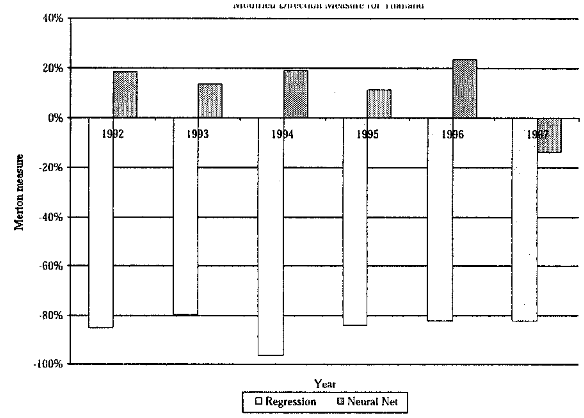

# EXHIBIT 23
## Modified Direction Measure for Mexico




# EXHIBIT 24
## Modified Direction Measure for Taiwan



## EXHIBIT 25
**Modified Direction Measure for Thailand**




that suffered through the Asian economic collapse of 1997-1998. Hence, now we now present the results from a similar analysis of Korea, using weekly data from June 1997 to March 1999. As before, the inputs to the net for training and the linear regression for significance testing were four lags of the return itself and one lag of the U.S. return. The investment strategies prescribed by the neural network and linear regression models and a buy-hold strategy are analyzed as before using return on $1, directional count, and the Merton Measure.

For these data with the above inputs, none of the linear regression models proved to be significant at the 90% confidence level that we had been using previously. Indeed, no matter what the input variables we used, the t-statistics were extremely low, that is, below 1. Hence, it must be stated that the linear regression did not produce a meaningful investment strategy for Korea over this period.

As shown in the Exhibits 26 and 27, the neural network prescribed a strategy that was invested in the Korean market for only twenty-five weeks during the ninety-four-week period during which the buy-hold strategy lost 35% of its value in very turbulent conditions. Thus, using the net strategy lowered the volatility of the portfolio and increased the return. The net strategy earned 79% on each dollar invested. The Merton Measure for the neural net was positive, albeit small, showing that the return given by the net's strategy was not merely due to a safe investment in risk-free Treasuries.

Hence this study in the Korean market during the economic crisis of 1997-1998 produces results consistent with (but more pronounced than) the complete analysis of the broader emerging market returns described in main section of this article.

### Limitations

All relationships are not linear. The benefit of the neural net is that it is able to discern non-linear relationships, and is therefore able to find predictive power in factors that may be useless in a simple linear model. However, one significant drawback is that the user is ultimately unaware of the closed-form relationship that the neural net has learned. Therefore, it is not possible to apply human intuition to the model regarding the sensibility of the modeled sensitivities. Such human intuition is often invaluable assessing the forward-looking viability of a particular relationship.

An additional drawback of the neural net is the model's excessive data requirements. Weekly data is difficult to acquire and potentially unreliable, particularly for emerging markets, which are often less efficient than the US market. For some time series, weekly data simply does not exist, or has not been collected for long period of time. For example, weekly data is not available for many economic time-series, since the statistics are only collected on a monthly or quarterly basis. This precludes the use of variables that might prove to offer explanatory power without interpolating between them, for example, by assuming they are constant, or in more generality linear, between data points.

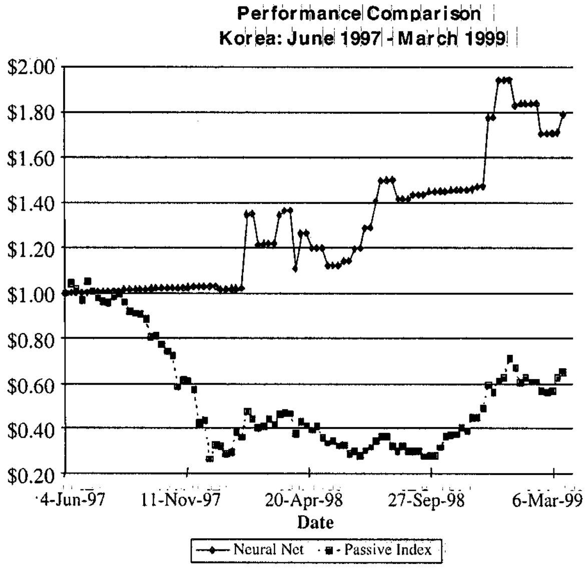
Further, the neural net's use of numerous data points lends itself to over-fitting of the model. This means that the neural network tries to incorporate back-

## EXHIBIT 26
**Performance of Neural Net for Korea During the Asian Crisis: June 1997 - March 1999**

| | Return | Correct Direction | Modified Direction | No. Down Weeks Invested | No. Up Weeks Invested |
|---|---|---|---|---|---|
| Neural Net | 79% | 59% | 16% | 9 | 16 |
| Index | (35%) | NA | NA | 46 | 48 |

# EXHIBIT 27

**PerformanceComparison Korea:June1997 – March 1999**




ground random noise in the in-sample data into the model. This obviously leads to inaccurate predications out-of-sample, where the noise is not predictable. We tried to minimize this effect by not using too many training runs on each sample. However, it is often difficult to find the optimal balance between sufficient training and over-fitting.

Finally, one important item that we have not considered is the high transactions costs of investing in emerging markets. Shifting weights between the individual emerging market and cash could induce prohibitive transactions costs that potentially could eliminate the return benefit of the net strategy versus the buy and hold. We leave the incorporation of transactions costs in the asset allocation strategy for further research.

## CONCLUSIONS

Previous research has suggested that significant predictability exists in emerging markets. However, the nature of the predictability changes through time. We test whether a non-linear modeling method, neural networks, has the ability to out perform standard benchmarks.

Our neural net outperformed both the active regression model and the passive buy-and-hold strategy over the 1992-1997 period. The process managed to "see through" the white noise created by the random elements of the financial markets, and identify the underlying nonlinear model sufficiently well to predict future returns in the emerging market indices.

Neural networks are still a topic of ongoing research, and therefore many of the analysis techniques may continue to be refined or altered in the future. At this point, there are very few significance tests analogous to those for linear regression techniques. Despite this, it was possible to identify qualitatively which inputs would be useful to train the model. However, due to the lack of quantitative statistical analysis, it is very important to ensure that the independent variables are economically reasonable so that modeling with them could be qualitatively justified.

Despite our success, it is important to be aware of the limitations of neural nets. During our research, we discovered that it was very easy to inadvertently over-fit the data by running the training for too many cycles using too many weeks of data. Great care needs to be exercised when implementing these models for active investing.

## REFERENCES


Bekaert, Geert, and Campbell R. Harvey. Time-Varying World Market Integration, Journal of Finance, 1995, pp. 403-444.

International Finance Corporation, Emerging Markets Data Base, January 1998.

International Finance Corporation, EMDB Guide for ASCII Users, January 1997.

Campbell, John Y., Andrew W. Lo, and Craig A. MacKinlay. *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Harvey, Campbell R. "The World Price of Covariance Risk," *Journal of Finance*, 46 (1991), pp. 111-157.

Harvey, Campbell R."Predictable Risk and Returns in Emerging Markets." *Review of Financial Studies*, 1995, pp. 773-816.

Merton, Robert. "On Market Timing and Investment Performance: An Equilibrium Theory of Value for Market Forecasts," *Journal of Business*, 1981.

Muller, B, J. Reinhardt, and M.T. Strickland. Neural Networks, an Introduction, 2nd ed., Springer-Verlag. 1995.

White, Halbert. *Artificial Neural Networks: Approximation and Learning Theory*, Cambridge, MA: Blackwell Publishers, 1992.