IPOX - An Initial Public Offerings IndeX

CHRISTIAN HÄFKE
CHRISTIAN HELMENSTEIN

Forschungsbericht/
Research Memorandum No. 351

September 1994

Institute for Advanced Studies
Stumpergasse 56
A-1060 Vienna, Austria
☎: (1) 59991-254
Fax: (1) 59991-163
E-mail: chris@ihssv.wsr.ac.at
E-mail: helmen@ihssv.wsr.ac.at

ABSTRACT:

In this paper we construct an index of Initial Public Offerings (IPOX) which is isomorph to the Austrian Traded Index (ATX). Conjecturing that the ATX constitutes a good leading indicator for the IPOX, in a first step we predict the ATX using linear and neural network models. Second, we estimate the IPOX one day ahead based on observed ATX data. We compare the quality of this estimation to an IPOX forecast based on forecasted ATX values in a third step. Given the ATX forecasting method the IPOX forecast based on a linear model always outperforms the ANN - IPOX forecast. By contrast, given the IPOX forecasting method, the ANN - ATX forecast beats the linear ATX - forecast.

ZUSAMMENFASSUNG:

In dieser Arbeit konstruieren wir einen zum Austrian Traded Index (ATX) isomorphen Erstemissionsindex (IPOX). Unter der Annahme, daß der ATX einen guten Leading Indicator für den IPOX darstellt, prognostizieren wir im ersten Schritt den ATX sowohl mit linearen Modellen als auch mit Neuronalen Netzen. In einem zweiten Schritt nehmen wir eine IPOX - Prognose für den Folgetag vor, die auf den realisierten ATX - Werten beruht. Wir vergleichen die Qualität dieser Schätzung mit einer IPOX - Prognose auf der Basis der prognostizierten ATX Werte in einem dritten Schritt. Unabhängig von der Prognosemethode für den ATX ergibt sich, daß das lineare IPOX - Schätzmodell bessere Ergebnisse liefert als das Neuronale Netz, die ATX - Prognose über das Neuronale Netz aber die Güte jeder IPOX - Prognose verbessert.

KEYWORDS: Artificial Neural Networks, Initial Public Offerings, Stock Price Indices

JEL-CLASSIFICATIONS: G14, G15, C22

The authors gratefully acknowledge constructive comments provided by Raimund Alt, Bernhard Felderer, Michael Hauser, Robert Kunst, and Andreas Wörgötter.

1. Introduction

Research in nonparametric, nonlinear methods has been extensive in recent years and three popular examples of these approaches are projection pursuit regression (FRIEDMAN and STUETZLE 1991), radial basis functions (e.g. POWELL 1987) and multilayer feedforward networks (RUMELHART et al. 1986). BARRON and BARRON (1988) called these three approaches, which can all be seen as nonparametric approaches to nonlinear regresion, learning networks. In this paper, however, we shall focus on multilayer feedforward networks. The two important areas of application of such networks are pattern classification and function approximation. BAXT and WHITE (1994) apply bootstrapping and neural networks to analyse medical data, MATTEWS and MERRIAM (1994) train a neural network on Shakespearean text and use it to determine whether "The Two Noble Kinsman" were really written by him or by John Fletcher. TESAURO reports of a neural net that learns to play backgammon (TESAURO 1989) and DIRST and WEIGEND (1994) analyse Bach's unfinished "*Die Kunst der Fuge*". GALLANT and WHITE (1988) show that *there exists a neural network that does not make avoidable mistakes*. They demonstrate how a certain class of neural networks possesses all properties of a Fourier series approximation. HORNIK, STINCHCOMBE and WHITE (1989, 1990) show that neural networks are universal approximators and can learn arbitrary functions arbitrarily well. A collection of the theoretical findings is presented by KUAN and WHITE (1994).

We apply a certain class of neural networks, an augmented single hidden layer feedforward network to Initial Public Offerings which have gained rather little attention in financial econometrics so far. In this paper we construct an index of Initial Public Offerings (IPOX) which is isomorph to the Austrian Traded Index (ATX). We conjecture that the ATX constitutes a good leading indicator for the IPOX. First, we predict the ATX using linear and neural network models. In a second step we estimate the IPOX one day ahead based on observed ATX data. We compare the quality of this estimation to an IPOX forecast based on forecasted ATX values in a third step. For all predicting purposes in this paper we estimate linear models as well as neural networks, as the latter are becoming increasingly common in financial forecasting as well (for example NNCM93, TRIPPI et TURBAN 1993). Section 2 analyses the data. Sections 3 and 4 present the models used. Section 5 discusses the error measures and the forecasting results. Section 6 contains some concluding remarks.

2. Data

2.1. The Austrian Traded Index

On the occasion of the foundation of the Austrian Futures and Options Exchange, ÖTOB, the Austrian Traded Index (ATX) was formed as a modern and reliable stock index for the Vienna Stock Exchange. The ATX serves as both a basis for futures and options contracts and as a market indicator which reflects a representative and liquid market segment of Austrian stocks with about 70% of total stock market activity (ÖTOB, 1991, 1994).

Three factors are essentially connected with the construction of a stock market index, viz. the selection of stocks, their weighting and the calculation method. The selection of stocks for the ATX follows the criteria of continuous trading, high market capitalisation and a sufficiently high free float. The weighting of a particular stock represents its equity market capitalisation. Thus, a title with a high equity market capitalisation (that is the number of stocks issued times the rate) has a larger impact on the index than a stock with a lower equity market capitalisation. The market capitalisation itself is corrected by a free float factor which ensures that the weight of a particular title in the ATX corresponds to the equity actually available for trading at the stock exchange.

The ATX is calculated according to the following formula

$$ATX_t = ATX_{t-1}\left[\frac{\sum_{i=1}^{n}\left(P_{i,t}Q_{i,(t-1)}\right)}{\sum_{i=1}^{n}\left(P_{i,(t-1)}Q_{i,(t-1)}\right)}\right], \tag{1}$$

with ATX_t denoting the ATX value at time t, $P_{i,\,t}$ denoting the price of share i at time t, $Q_{i,(t-1)}$ the number of shares of stock i issued (corrected by the free float factor) and n the number of stocks in the ATX.

While the ATX reflects all price changes due to market fluctuations, technical price changes do not affect the index. For this reason the prices of the underlying shares are adjusted to changes in the capital of a title or dividend payments. The basis of the ATX is 1000 as per January 2^{nd}, 1991, the ATX value as of April 27^{th}, 1994 is 1089.42. Table A1 shows the composition of the ATX as of the latter date. The data used for estimation is of length 1691 and ends on November 1^{st}, 1992. The consecutive 345 observations were used as test set for the out of sample error measures.

2.2. The Initial Public Offerings IndeX

A distinctive feature of an Austrian IPO is the prospectus which provides investors with more information than they would have at their disposal in the case of an ordinary share. Of particular interest are the projections of expected future profits by the company itself. Furthermore, the bank(s) underwriting the issue can, besides others, be held liable for wrong or misleading statements in the prospectus. Thus, the prospectus contains relatively more comprehensive and reliable information than any other information source available to the outside investor. This feature of an Austrian IPO should, at least, be reflected in two market statistics. First, the average IPO should exhibit a lower volatility as compared to the market average since the discounted value of future profits is less uncertain. Second, with investors' extended information the risk premium should diminish, and hence, the yield of the share should be lower.

In order to render the ATX and the IPOX comparable to each other and to exclude a systematic deviation of the IPOX from the ATX, the IPOX is constructed isomorphically to the ATX. The IPOX covers all initial public offerings in the official market segment, while initial public offerings in the regulated and unregulated market segments are excluded from consideration. Each individual initial public offering enters the IPOX with the first rate in public trading and not with the offering price.

If we consider the additional information as being a typical attribute of a share to be defined as an IPO, we have reason to expect that this status will vanish at the end of the forecasting horizon in the prospectus which is one and a half years on average. For this reason one and a half years past the first listing on the stock exchange a stock does no longer qualify as an IPO and is withdrawn from the index. Subsequently, we will also disregard the three different categories of Bank Austria AG stocks. As opposed to the other banks in the index, Bank Austria AG holds a significant amount of shares of several other Austrian stock companies. Since our analysis is concerned with the particular portfolio of Austrian IPOs, we exclude a title which might be regarded as a particular portfolio of Austrian stock companies itself for the sake of analytical clarity. This argument is valid a fortiori as the last but one column of table A2 illustrates that Bank Austria AG stocks account for 48.52% of the total capital covered by the IPOX at the beginning of our sample period, November 1992. At other points of time its share would be even higher.

For estimating the parameters of our models we use the first 276 observations beginning on November 2nd, 1992. The remaining 69 observations are used to calculate the out of sample error measures.

2.3. Autocorrelation and Cross-correlation of ATX and IPOX

After taking the logarithms of the ATX and the IPOX we compute sample autocorrelations of their first differences. The results replicate the findings of PICHLER (1993) that Austrian time series data of indices exhibits significant sample autocorrelations of order 1 (table1). Our analysis confirms this outcome for another subsample of Austrian stocks, the initial public offerings.

Table 1: Autocorrelations for observations 1 to 345

Variable	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7
IPOX	0.2248	0.0958	0.0359	0.0620	0.0010	-0.00004	0.0400
ATX	0.2410	0.0130	-0.1141	0.0106	0.1192	0.0611	-0.0230

As the ATX covers the most liquid shares at the Vienna Stock Exchange, it might qualify as a leading indicator for the IPOX. Support for this hypothesis comes from the computation of cross-correlations between the ATX and the IPOX (table 2).

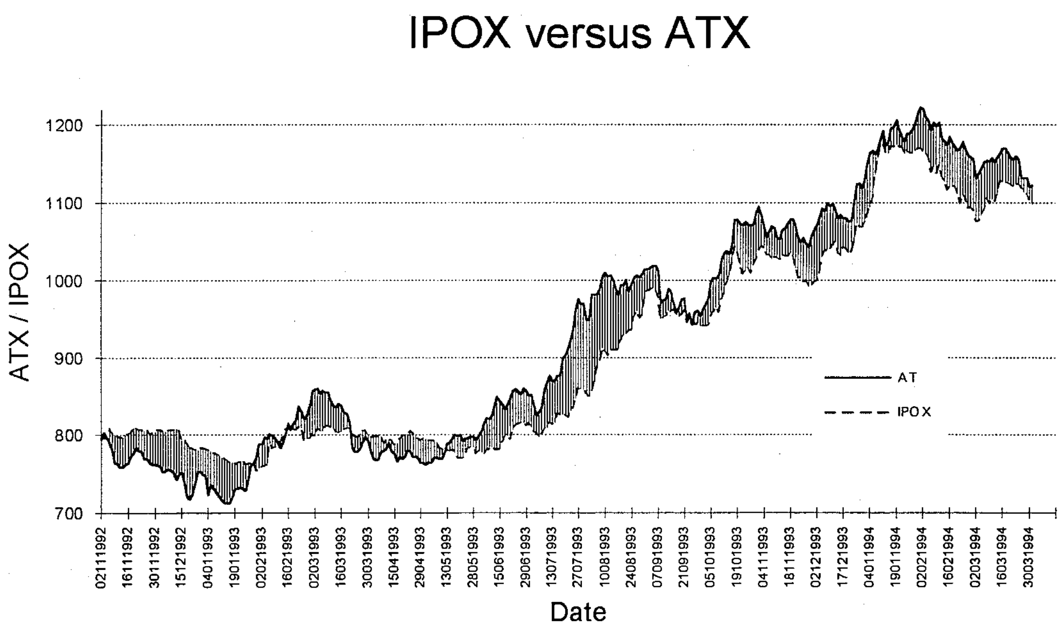

Figure 1: IPOX versus ATX

Table 2: Cross - correlations between the ATX and the IPOX for observations 1 to 345

Lags	Lag -3	Lag -2	Lag -1	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5
X - correlation	-0.0681	0.0281	0.0673	0.4760	0.2074	0.0910	0.0020	0.1025	0.0723

According to these results the ATX tends to be one period ahead of the IPOX. These findings should subsequently be confirmed by ordinary least squares regressions.

3. Linear Model

The forecast of the IPOX[1] is based on a dynamic specification of the linear regression model

$$dIPOX_t = \alpha_0 + \alpha_1 dIPOX_{(t-1)} + \alpha_2 dATX_{(t-1)} + \alpha_3 (IPOX - ATX)_{(t-1)} + \varepsilon_t. \quad (2)$$

We use the same explaining variables to model dATX in order to check the above hypothesis,

$$dATX_t = \beta_0 + \beta_1 dIPOX_{(t-1)} + \beta_2 dATX_{(t-1)} + \beta_3 (IPOX - ATX)_{(t-1)} + \eta_t. \quad (3)$$

The regression results give some evidence that the current value of the IPOX is positively related to the previous values of the IPOX and the ATX. The highly significant value for the error correction term (IPOX-ATX) with a lag of order 1 reveals that deviations of the IPOX from the ATX cause a strong pull back tendency towards the ATX, while the opposite does not hold. The insignificant result for a dependency of the ATX on the IPOX additionally strenghtens these findings.

[1] In the linear context IPOX and ATX always denote the logarithm of the respective series.

Table 3: The ATX as a leading indicator for the IPOX

Explaining variable	Dependent variable	
	$dIPOX_t$	$dATX_t$
Intercept	-0.0000064 (-0.0153)	0.000682 (1.1681)
$dIPOX_{t-1}$	0.1447** (2.4944)	-0.0893 (-1.1014)
$dATX_{t-1}$	0.0756* (1.7585)	0.2663** (4.4334)
$IPOX\text{-}ATX_{t-1}$	-0.0311** (-3.6588)	-0.00694 (-0.5834)
t-statistics are given in parentheses. * Statistically significantly different from zero at the 0.05 significance level. ** Statistically significantly different from zero at the 0.01 significance level.	$\bar{R}^2$ = 0.0945 DW = 1.9874 LM§: p - value = 0.8285 Ljung-Box Q (54) = 57.4 (p - value = 0.3499)	$\bar{R}^2$ = 0.0545 DW = 1.9734 LM§: p - value = 0.2568 Ljung-Box Q (54) = 51.4 (p - value = 0.5753)

The final linear model chosen for the dATX forecast is specified as an AR[2] process, that is

$$dATX_t = \gamma_1 dATX_{(t-1)} + \gamma_2 dATX_{(t-2)} + \varepsilon_t \tag{4}$$

A constant γ_0 was used during estimation but turned out to be insignificant and is therefore not taken into consideration for the forecasting procedure.

Table 4: Estimated coefficients for the linear model

Explaining variables	dATX
$dATX_{(t-1)}$	0.3568** (14.7062)
$dATX_{(t-2)}$	-0.0843** (-3.4730)

§ The LM test is performed to test for first order autocorrelation.

4. Neural Network Model

The idea of modelling brain functionality goes back to McCULLOCH and PITTS (1943) who first introduced units that were constructed analogously to a neuron in the brain. By combining these neurons we arrive at a linear perceptron, or ADALINE, as it was first called by WIDROW and HOFF (1960).

$$f(\tilde{x}_t, \alpha) = G(\tilde{x}_t'\alpha) \tag{5}$$

with $\tilde{x}_t$ being the input vector augmented by a constant and α a set of weights. By taking a closer look at the formula, we see that for G(x) = x we arrive at the simple linear model which is a standard paradigm in economic and econometric modelling. KUAN and WHITE (1994) point out, that for G(x) = 1/(1+exp(x)) we arrive at the binary logit model and for G(x) being any normal cumulative distribution function we obtain a binary probit. We see that even at the outset of neural network modelling, standard econometric models could easily be included as special cases.

With all these possibilities, the Perceptrons' popularity kept increasing until MINSKY and PAPERT (1969) published their book Perceptrons, where they pointed out that Perceptrons were only capable of solving linearly separable problems. This alone could not have been the reason why research in neural networks consequently diminished rapidly as there exist methods of transforming any problem into a linearly separable problem.

Another way to overcome the separability problem was obtained by looking at nature again. It is very rare that signals directly flow from the sending to the receiving cell. They usually pass a number of intermediate layers. These intermediate layers were also adopted by the neural network community (ROSENBLATT, 1958). However, until WERBOS (1974) and RUMELHART (1986) there existed no way of actually estimating such networks. The output produced by such a multi layer perceptron is given by:

$$f(\tilde{x}_t, \beta, \gamma) = F\left(\sum_{j=1}^{q} G(\tilde{x}_t'\gamma_t)\beta_t\right) \tag{6}$$

Standard neural network *learning* algorithms use incremental updates of the form

$$\hat{\theta}_{(t+1)} = \hat{\theta}_t + \eta \nabla f\left(\tilde{x}_t, \hat{\theta}_t\right)\left(y_t - f\left(\tilde{x}_t, \hat{\theta}_t\right)\right) \tag{7}$$

with $\tilde{x}$ denoting the input vector x augmented by a constant, and θ denoting a weight vector. WHITE (1987) pointed out that this is actually a form of stochastic approximation (ROBBINS and MONRO, 1951) where η is fixed over time instead of dependant on t. For further discussion see KUAN and WHITE (1994). There are now a number of approaches that explicitly allow for a time varying learning rate (η). It has been useful to start with a high η and slowly decrease it, which incorporates a special form of simulated annealing.

Many different forms of neural networks are successfully applied to time series data with the simple single hidden layer network being one of them (for example TAN 1993, LEE and PARK 1992, LI et al. 1990, WHITE 1988). However, it has frequently been noted that performance sometimes degrades after adding a hidden layer as compared to a simple perceptron. To avoid these shortcomings we use an augmented single hidden layer feedforward neural network as proposed for example by SWANSON and WHITE (1992) which incorporates all the models discussed so far and thus constitutes a very flexible model for econometric tasks. This structure incorporates a simple perceptron and a simple single hidden layer network. Therefore the output is calculated as follows:

$$f(\tilde{x}_t, \theta) = \tilde{x}_t'\alpha + \sum_{j=1}^{q} G(\tilde{x}_t'\gamma_t)\beta_t \tag{8}$$

with $\tilde{x}$ denoting the input vector x augmented by a constant, and θ denoting a weight vector containing the weights α, β, γ, that is $\theta = (\alpha', \beta', \gamma')'$, $\beta=(\beta_1, \beta_2, \ldots\beta_q)'$, $\gamma=(\gamma_1', \ldots \gamma_q')'$. q is the number of Hidden Units and G is a nonlinear function, in this case

$$G(x) = \frac{2}{1+e^{-x}} - 1, \tag{9}$$

thus mapping x into the [-1;+1] interval.

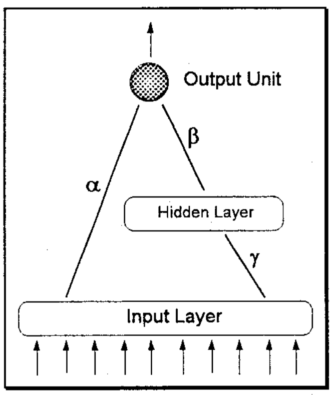

Figure 2: The Neural Net Model

This architecture not only allows us to capture the nonlinearity in the data but also makes use of the well known linear regression approach and therefore ensures that in sample the ANN will perform at least as good as a linear model. If the input - output connections were dropped, this could not be guaranteed.

Training actually takes place in two steps. First, the direct input-output connections α are estimated through OLS and fixed. In a first step we estimate

$$f(x_t, \alpha) = \tilde{x}_t'\alpha + \varepsilon_t, \tag{10}$$

with $\tilde{x}_t$ being the input vector x augmented by a constant, α the corresponding weight vector, and ε_t the vector of the residuals. Matrices β and γ are estimated to model the residuals of the linear regression with any nonlinear optimisation technique. This approach generally improves performance over OLS. In a second step we solve the problem:

$$\min_{\theta} \frac{1}{n}\sum_{i=1}^{n}\left(y_t - f(\tilde{x}_t, \theta)\right)^2 \tag{11}$$

with α fixed. Our programme is designed to find the optimal number of hidden units itself, using out of sample MSE as suggested in WHITE (1990). This approach helps increasing the generalisation performance of the net, which is a topic of keen interest (for example MOODY 1992, TERÄSVIRTA 1993, MOODY and UTANS 1994).

5. Empirical Results

The quality of our results is evaluated using the following out of sample error measures:

- out of sample MSE
- out of sample R^2

$$R^2 = 1 - \frac{\sum (y_t - \hat{y}_t)^2}{n \cdot S_y^2} \tag{12}$$

  n is the number of out of sample observations, S_y^2 the out of sample variance of the dependent variable (SWANSON and WHITE, 1992);

- Theil's measure of inequality

$$Theil = \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - y_{t-1})^2} \tag{13}$$

  This measure constitutes a simple sanity check of our forecasts against a no-change forecast, which performs better for $Theil > 1$ (THEIL 1966);

- Normalised Mean Squared Error

$$NMSE = \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y})^2} \tag{14}$$

  NMSE is a second sanity check against the out of sample mean of the dependent variable. This measure is used by WEIGEND and GERSHENFELD (1994) to evaluate entries into the Santa Fe Time Series Competition;

- Confusion Matrix

  The forecasts obtained through a feedforward pass through the network are then evaluated, and the up and down signals of the net are used to compute a confusion matrix as in SWANSON and WHITE (1992). We find the number of correct classifications in the main diagonal and the errors off the diagonal. The confusion rate is calculated as the sum of the off diagonal elements divided by the total number of elements. A binomial test is performed to check if the number of correct classifications differs significantly from 50%.

act / forec	Up	Down
Up		
Down		

Figure 3: Confusion Matrix

- Trading Scheme

We apply a very simple and conservative trading scheme without transaction costs. We start out on the first day of our evaluation period. If the forecast for the following day indicates a rise in prices and we do not yet hold the IPOX - portfolio we buy. In case we already hold it we do not buy more but just keep still. In case of falling prices we sell if we hold but never go short. Returns are annualised and compared to a Buy and Hold strategy and to a case of perfect foresight, which thus represents the maximum return achievable through the strategy.

- Moving Average Trading Rule

We also compare our returns against the returns generated by a 2-50 MA-Trading Rule. If the short MA intersects the long MA from below we receive a buy signal and keep the stock until the two moving averages intersect again and vice versa.

- t - values for returns of the Trading Scheme

t - values were computed to test whether the returns generated through the trading scheme are significantly different from the Buy and Hold strategy according to the following formula (BROCK, LAKONISHOK, LEBARON 1991):

$$t = \frac{\mu_t - \mu_b}{\sqrt{\frac{\sigma^2}{N_t} + \frac{\sigma^2}{N}}} \tag{15}$$

with μ being the mean returns of the two series, σ^2 the estimated variance for the entire sample, N_t the number of trades and N the number of observations. It is worth noting that models with low Theil and NMSE are not the best ones for forecasting the sign of the index movement. However, those with good confusion rates usually perform worse than no - change forecasts with respect to the MSE. Therefore the choice of the correct model strongly depends on the aim of the forecast. In our case we try to approximate the movements of the ATX and IPOX as good as possible and therefore select the model according to MSE criteria. However, for a trading scheme based on the sign of the change it would be much more sensible to disregard the size but focus on the sign, that is the confusion rate.

For the ATX the following results are achieved:

Table 5: Results of out of sample ATX forecasts

Error measures	Linear Model	Best ANN
MSE	231.793	85.145
R^2	0.9907	0.9966
Theil	2.631	0.966
NMSE	0.00933	0.00343
Confusion Matrix[*]	$\begin{bmatrix} 119 & 69 \\ 74 & 83 \end{bmatrix}$ (3.22)	$\begin{bmatrix} 107 & 74 \\ 86 & 78 \end{bmatrix}$ (1.35)

We see that the nonlinear model with 6 lags and 2 hidden units clearly outperforms the AR[2] with regard to the MSE criteria. However, when we look at the Theil measure, we still detect a great potential for improving the forecast which will be left for further work.

The linear IPOX reports the following results:

Table 6: Results of out of sample IPOX forecasts with a linear model

	ATX forecasts generated through		
Error measures	Linear Model	Best ANN	Observed ATX
MSE	71.698	67.583	61.797
R^2	0.945	0.948	0.953
Theil	0.846	0.797	0.730
NMSE	0.055	0.052	0.048
Confusion Matrix[*]	$\begin{bmatrix} 17 & 13 \\ 16 & 23 \end{bmatrix}$ (1.34)	$\begin{bmatrix} 15 & 16 \\ 18 & 20 \end{bmatrix}$ (0.12)	$\begin{bmatrix} 19 & 13 \\ 14 & 23 \end{bmatrix}$ (1.85)

We compare these findings to the results obtained through an estimation of the IPOX with an

[*] t - values for the binomial test, whether correct classifications are statistically greater than 50%.

[*] t - values for the binomial test, whether correct classifications are statistically greater than 50%.

artificial neural network with 1 Hidden Unit:

Table 7: Results of out of sample IPOX forecasts with ANN model

Error measures	ATX forecasts generated through: Linear Model	ATX forecasts generated through: Best ANN	Observed ATX
MSE	86.830	79.888	77.272
R^2	0.933	0.939	0.941
Theil	1.024	0.943	0.912
NMSE	0.067	0.062	0.060
Confusion Matrix[*]	$\begin{bmatrix} 17 & 17 \\ 16 & 19 \end{bmatrix}$ (0.36)	$\begin{bmatrix} 16 & 19 \\ 17 & 17 \end{bmatrix}$ (-0.36)	$\begin{bmatrix} 17 & 16 \\ 16 & 20 \end{bmatrix}$ (0.60)

It turns out that, given the ATX forecasting method, the IPOX forecast based on a linear model always outperforms the ANN - IPOX forecast. By contrast, given the IPOX forecasting method, the ANN - ATX forecast always beats the linear ATX - forecast.

We see that the ANN does not boost the performance of the IPOX forecasts. The ATX *forecasts*, however, seem to contain some information which is absent in the linear ATX forecasts. Actually, our IPOX models are estimated as nested models where we start out from a linear model and add Hidden Units to account for nonlinearity. However, it seems that our model selection criterion, namely hold out crossvalidation did not properly account for the high number of parameters in Neural Networks and was biased towards too large models. Similar problems were encountered by SWANSON and WHITE (1992) when they regarded SIC (SCHWARTZ 1978).

[*] t - values for the binomial test, whether correct classifications are statistically greater than 50%.

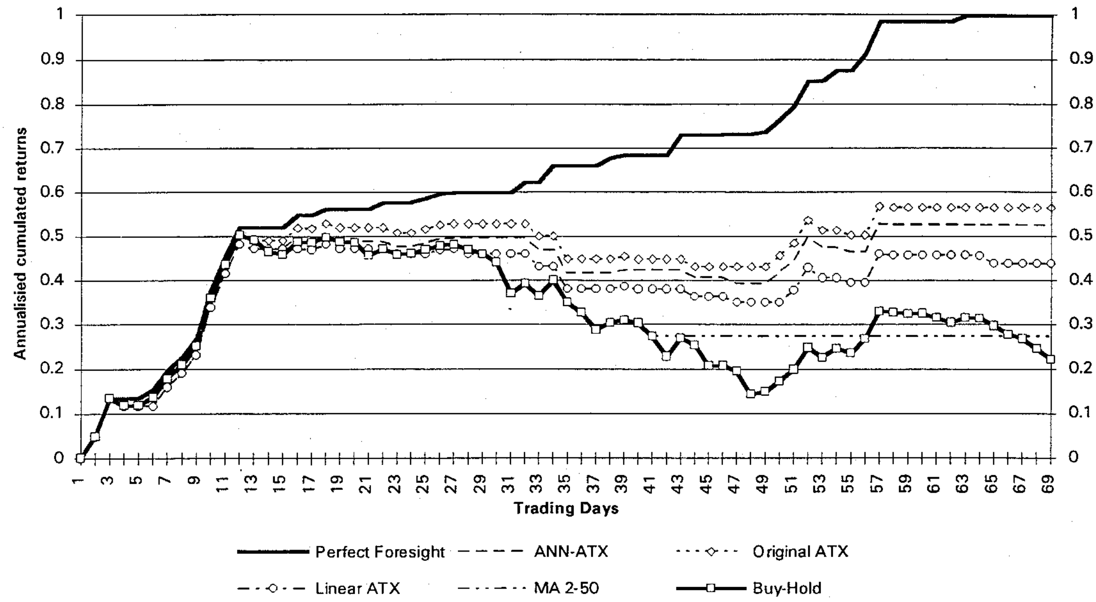

Figure 4: Annualised Cumulated Returns for the Linear IPOX estimation

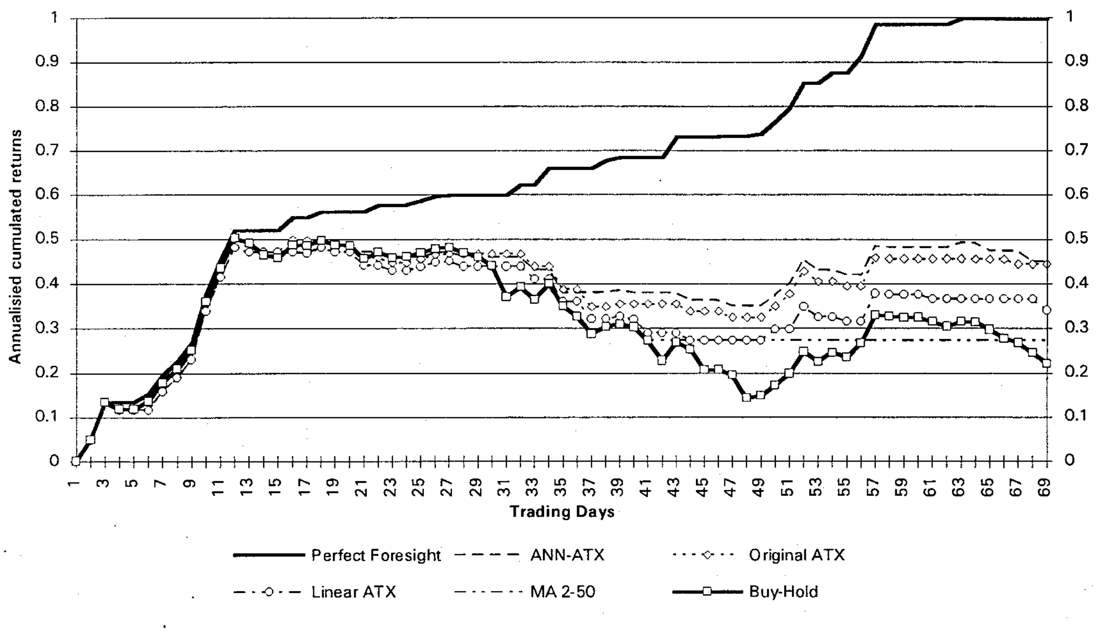

Figure 5: Annualised Cumulated Returns for the ANN IPOX estimation

Table 8: Summary statistics for returns of various forecasts

	cumulated returns	Number of transactions	t-value (vs. Buy Hold)
linear IPOX linear ATX	0.438	28	4.63
linear IPOX ANN ATX	0.526	30	4.86
linear IPOX orig ATX	0.565	26	4.72
ANN IPOX linear ATX	0.341	35	5.00
ANN IPOX ANN ATX	0.451	30	4.86
ANN IPOX orig ATX	0.445	36	5.29
MA 2-50	0.274	22	0.59

The analysis of the trading schemes yields the following results. As benchmarks we use a simple buy and hold strategy, the 2-50 Moving Average and - in order to determine the highest possible return - a strategy based on perfect foresight of the next period's IPOX value. First, we consider the linear approach to forecast the IPOX. We compare three datasets. Two of the underlying ATX series are generated by a linear model and a neural network, the third one is made up by the original ATX series. The ANN dataset is very close to the original dataset but the returns of the linear approach fall remarkably short of the other two. However, all of them significantly beat the buy and hold strategy as well as the moving average. Second, modelling the IPOX with an ANN, we achieve similar results. The original data series and the ANN ATX based IPOX estimation are quite close again and far better than the returns based on the linear ATX. Compared to the first case the profitability of all forecasts is lower.

6. Conclusions

The ATX forecast using ANN yields remarkably better results relative to the linear model. Particularly the Theil measure demonstrates that the linear model performs by far worse than a no - change forecast whereas the ANN performs slightly better than the naive approach. Nevertheless, the ATX forecast remains to be improved in order to fully exploit the superior forecasting capabilities of Artificial Neural Networks in connection with the IPOX forecast. The property of the ATX to be a leading indicator for the IPOX should be exploitable by any market participant capable of forecasting the ATX one day ahead. Therefore it would be of considerable interest to investigate the profitability of such models applying various trading schemes. In particular it remains to be inquired whether these profits sustain in an environment where transaction costs are accounted for if we consider that we traded between 22 and 36 times (depending on the forecast) in just 70 days. Furthermore, other error functions for the neural network should be considered, the first is minimising the confusion rate, the other maximising profits obtained through transactions based on the above trading rules. As we noted above the use of more conservative model selection criteria might help increase the generalisation performance of the networks, and thus other estimators for the prediction risk shall also be taken into consideration. With all these amendments we expect to extract even more information from the ATX - IPOX relation.

Appendix

Table A1: ATX - stocks

Stock	Number of shares	Price	Free float factor	Weight in the ATX (April 27th, 1994)
Austrian Airlines AG	2600000	1940.00	0.5	0.021671757
Bank Austria AG Vz.	6826500	619.00	1	0.036310964
Bundesländer Vers.-AG Vz.	725000	672.00	1	0.00418655
Constantia Industrieholding AG	3520000	1010.00	0.5	0.015275065
Creditanstalt-Bankverein Vz.	12300000	680.00	1	0.071872552
EA Generali AG St.	7000000	3740.00	0.5	0.112483465
EVN AG	9500000	1450.00	1	0.118369727
Flughafen Wien AG	5400000	488.00	1	0.022644494
Lenzing AG	3675000	1289.00	1	0.040706082
Leykam-Mürztaler AG	9803459	425.00	0.5	0.017901413
Maculan Holding AG Vz.	1139626	1113.00	1	0.010899499
ÖMV AG	24000000	966.00	1	0.199222051
Öster. Elektrizitätswirtschafts AG	15101800	631.00	1	0.08188552
Radex-Heraklith Industriebet.-AG	6928779	460.00	1	0.027388206
Steyr-Daimler-Puch AG	10000000	237.00	1	0.020365608
Strabag Österreich AG	1440000	1789.00	1	0.022137158
Universale-Bau AG	3000000	855.00	1	0.02204126
Veitsch Radex AG	9809125	346.00	1	0.029164559
Wienerberger Baustoffindustrie AG	3935782	3710.00	1	0.12547407

St.: common stock; Vz.: preferred stock.
Without special notice: only one category of equity existent (common stock).

Table A2: IPOX stocks

Stock	Capital (in mill. AS)	Price (adjusted)	Weight in the IPOX (Nov 30th, 93)	(Nov 2nd, 92)	(Nov 2nd, 92)
Agrana Vz.	150	292.00		0.007721966	0.015000375
Bank Austria St.	1,883	1036.00		0.343925061	n.a.
Bank Austria Vz.	683	612.00		0.073692938	n.a.
Bank Austria PS	1,009	380.00		0.067597169	n.a.
Schärdinger St.	67	1140.00		0.013465839	0.026158188
Voith Vz.	28	250.00		0.001234104	0.00239732
IGM St.	50	620.00		0.005465318	0.010616704
Frauenthal St.	69	720.00		0.008758613	0.017014124
ÖMAG St.	650	285.00		0.032659683	0.063443367
Wienerberger Immob. St.	800	1930.00		0.2722081	0.528780343
Viso-Data St.	60	1940.00		0.020521388	0.03986401
UBM Vz.	15	1005.00		0.002657731	0.0051628
Oberbank Vz.	30	596.00		0.003152254	0.006123441
Kies-Union St.-NA	40	2940.00		0.020732949	0.040274979
BKS Vz.	12	504.00		0.001066266	0.002071285
BWT St.	150	1755.00		0.046411128	0.090156363
Flughafen Wien St.	540	517.00	0.070557982	0.049219597	0.095611979
Kapital & Wert St.	252	442.00	0.028150406	0.019637064	0.038146159
Binder & Co St.	50	1120.00	0.014153045	0.009872833	0.019178562
VAE St.	120	1700.00	0.05155752		
Pengg Kabel St.	60	1100.00	0.016680374		
Investitionskredit St.	500	4650.00	0.587604086		
AMS St.	250	619.00	0.039110423		
Bau Holding St.	425	1220.00	0.13104203		
Erste Vz.	436	540.00	0.059503443		
BTV Vz.	10	649.18	0.001640692		
Σ			1	1	1

PS: participation certificate.

References

BARRON, A.R. and BARRON, R.L. (1988): Statistical Learning Networks: a unifying view, *20th Symposium on the Interface: Computing Science and Statistics*, pp 192-203.

BAXT, W. and WHITE, H. (1994): Bootstrapping Confidence Intervals for Clinical Input Variable Effects in a Network Trained to Identify the Presence of Acute Myocardial Infarction, *Technical Report,* University of California, San Diego.

BROCK, W., LAKONISHOK, J., LeBARON, B. (1991): Simple Technical Trading Rules and the Stochastic Properties of Stock Returns, *Journal of Finance*, vol 47, pp 1731-1764.

DIRST, M. and WEIGEND, A.S. (1994): Baroque Forecasting: On Completing J. S. Bach's Last Fugue, in WEIGEND, A.S. and GERSHENFELD, N.A. eds., *Time Series Prediction Forecasting the Future and Understanding the Past,* Addison Wesley: Reading, MA, pp 151-172.

FRIEDMAN, J.H. and STUETZLE, W. (1981): Projection Pursuit Regression, *Journal of the American Statistical Association,* vol 76, pp 817-823.

GALLANT, A. R. and WHITE, H. (1988): There Exists a Neural Network that Does Not Make Avoidable Mistakes, *Proceedings of the Second Annual IEEE Conference on Neural Networks*, San Diego, IEEE Press: New York, vol 1, pp 657-664.

HORNIK, K., STINCHCOMBE, M., and WHITE, H. (1989): Multi-Layer Feedforward Networks Are Universal Approximators, *Neural Networks*, vol 2, pp 359-366.

HORNIK, K., STINCHCOMBE, M., and WHITE, H. (1990): Universal Approximation of an Unknown Mapping and Its Derivatives, *Neural Networks*, vol 3, pp 551-560.

KUAN, C., and WHITE, H. (1993): Artificial Neural Networks: An econometric perspective, *Econometric Reviews,* vol 13, N°1, pp 1-91.

LEE, K., and PARK, J.(1992): Short-term load forecasting using an artificial neural network, *IEEE Transactions on Power Systems*, vol 7, N°1, pp 124-132.

MATTHEWS R., MERRIAM, T (1994): in *Literary and Linguistic Computing, forthcoming*, taken from *N*ew Scientist, January 22nd, 1994, p 23.

McCULLOCH, W.S., and PITTS, W. (1943): A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics*, vol 5, pp 115-133.

MOODY, J. (1992): The *Effective* Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems, *Advances in Neural Information Processing Systems 4*, Morgan Kauffmann Publishers: San Mateo, CA, 1992.

MOODY, J., UTANS, J. (1994): Architecture Selection Strategies for Neural Networks: Application to Corporate Bond Rating Prediction, in: REFENES, A.N.: *Neural Networks in the Capital Market*, John Wiley & Sons: New York, forthcoming.

NNCM93 (1993): Neural Networks in the Capital Markets, Conference at the London Business School, November 1993.

ÖTOB (1994, 1991): ATX Austrian Traded Index, ÖTOB Clearing Bank AG: Vienna.

PICHLER, S. (1993): Size Effect und Settlement Effect am österreichischen Aktienmarkt (Size Effect and Settlement Effect on the Austrian Stock Market), *Österreichisches BankArchiv*, vol 3, pp 195-201.

POWELL, M.J.D. (1987): Radial Basis Functions for Multivariable Interpolation: A Review, in J.C. MASON and M.G. COX, eds., *Algorithms for Approximation,* Clarendon Press: Oxford, pp 143-167.

ROBBINS, H. and MONRO, S. (1951): A Stochastic Approximation Method, *Annals of Mathematical Statistics,* vol 22, pp 400-407.

ROSENBLATT, R. (1958): The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psychological Reviews*, vol 62, 386-408.

RUMELHART, D.E., HINTON, G.E., WILLIAMS, R.J. (1986): Learning Internal Representations by Error Propagation, in D:E. Rumelhart and J.L. McClelland eds., *Parallel Distributed Processing: Explorations in the Mircostructures of Cognition*, MIT Press: Cambridge, vol 1, pp 318-362.

SCHWARTZ, G. (1978): Estimating the Dimension of a Model, *Annals of Statistics*, vol 6, pp 461-464.

SWANSON, N., and WHITE, H. (1992): A Model Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks, Research Group for Econometric Analysis, Department of Economics, University of California, San Diego.

TAN, C. (1993): Hypothetical Trading Result with a Simple Artificial Neural Network-based Financial Trading System, paper presented at AIEM3, Portland, Oregon.

TERÄSVIRTA, T., and LIN, C. (1993): Determining the number of hidden units in a single hidden-layer neural network model, paper presented at the *NBER Time Series Seminar*, October 21, Institute for Advanced Studies, Vienna.

TESAURO, G. (1989): Neurogammon Wins Computer Olympiad, *Neural Computation*, vol 1, pp 321-323.

THEIL, H. (1966): Applied Economic Forecasting. North Holland Publishing Company: Amsterdam.

TRIPPI, R., and TURBAN, E. (eds.) (1993): Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance, Probus Pub. Co: Chicago, Ill.

WEIGEND, A.S., and GERSHENFELD, N.A. (eds.) (1994): Time Series Prediction, Forecasting the Future and Understanding the Past, Proceedings in the SFI studies of complexity, Addison-Wesley, Reading: MA.

WERBOS, P. (1974): Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, *unpublished Ph.D. Dissertation,* Harvard University, Department of Applied Mathematics.

WHITE, H. (1987): Some Asymptotic Results for Back-Propagation, *Proceedings of the IEEE First International Conference on Neural Networks,* San Diego, IEEE Press: New York, vol 3, pp 261-266.

WHITE, H. (1988): Economic prediction using neural networks: the case of IBM daily stock returns, *Proceedings of the Second Annual IEEE Conference in Neural Networks*, San Diego, IEEE Press: New York, vol 2, pp 451-458.

WHITE, H. (1989): An Additional Hidden Unit Test for Neglected Nonlinearity, *Proceedings of the International Joint Conference on Neural Networks*, Washington D.C. IEEE Press: New York, vol 2, pp 451-455.

WHITE, H. (1990): Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings, *Neural Networks*, vol 3, pp 535-549.

WIDROW, B. and HOFF, M.E. (1960): Adaptive Switching Circuits, *Institute of Radio Engineers WESCON Convention Record*, Part 4, pp 96-104.