ARTIFICIAL NEURAL NETWORKS FOR VALUATION OF FINANCIAL DERIVATIVES AND CUSTOMIZED OPTION EMBEDDED CONTRACTS

by
Christakis Charalambous and Spiros H. Martzoukos

Working Paper 01-19

HERMES Center of Excellence
on Computational Finance & Economics
University of Cyprus
P.O. Box 20537, 1678 Nicosia, CYPRUS

HERMES Center of Excellence on Computational Finance & Economics

The HERMES Center on Computational Finance and Economics at the University of Cyprus has been selected by the European Commission as a European Center of Excellence in 2000. The Center faculty, graduate students, and visitors pursue a broad research agenda that focuses on optimal financial decision making from both the supply side (financial institutions) and the demand side (households and institutional investors). Emphasis is placed on the challenges created for both sides by the globalization and innovations of the financial markets, especially for the economies of pre-accession States as they move towards harmonization with European Union.

The work of the Center is divided in five major areas. The first deals with enterprise wide risk management and the development of both innovative methodologies and decision support tools. The second deals with the saving and borrowing behavior of households and their portfolio choices. The third deals with empirical studies of capital markets and the information they reveal for the prediction of bankruptcy events and the management of credit risk. The fourth deals with real options, their valuation, and their use in coping with an uncertain world. The fifth deals with issues surrounding the performance of financial institutions in delivering quality services.

Data-driven theoretical modeling, empirical analysis, and the use of computations are the cornerstones of the disciplinary approaches of the Center's activity, that brings together researchers from finance, economics, accounting, management sciences, mathematics and computer science. Research is basic, yet relevant to this important sector of economic activity.

The Center fosters a community of scholars, practitioners, and graduate students whose interests support the mission of the Center. Close collaboration with industry ensures that the Center's research remains not only cutting-edge in pursuit of academic excellence, but is also relevant to financial institutions, in their quest for competitive excellence. If you would like to know more about the Center or join in its activities please let us know of your interest and visit our Web site at http://www.hermes.ucy.ac.cy/

Stavros A. Zenios
Director

Artificial Neural Networks for Valuation of Financial Derivatives and Customized Option Embedded Contracts

Christakis Charalambous and Spiros H. Martzoukos

Department of Public and Business Administration,
and HERMES Center of Excellence on Computational Finance and Economics,
University of Cyprus

This version, May 2001

Key words: Artificial neural networks, option pricing, investment contracts with guarantees, venture capital, and real options.

Comments by participants at the Asset and Liability Management Conference (Nicosia, May 2001) and partial support by the Hermes Center of Excellence on Computational Finance and Economics at the University of Cyprus (EC grant ICA1-CT-2000-70015) are appreciated.

Address correspondence:

Christakis Charalambous, Professor of Management Science,
or Spiros H. Martzoukos, Assistant Professor of Finance,
University of Cyprus, Department of Public and Business Administration,
75 Kallipoleos Str., P.O.Box 20537, CY 1678 Nicosia – Cyprus.
Department Tel.: 357-2-892461/2, Fax: 357-2-892460.
Email: baspiros@ucy.ac.cy and bachris@ucy.ac.cy

Artificial Neural Networks for Valuation of Financial Derivatives and Customized Option Embedded Contracts

Abstract

In this paper we propose and test a valuation methodology for improving the efficiency of contingent claims pricing using Artificial Neural Networks (ANN). Contingent claims is by now a standard method for pricing under uncertainty non-linear (option embedded) contracts, for both financial options (standardized or customized) and real (investment) opportunities. In the presence of liquid option markets, implied volatility surfaces have aided considerably option pricing with and without the use of ANN. In the absence of such liquid markets, customized positions are much harder to evaluate. The method in this paper improves the efficiency of valuation and financial decision-making dramatically. The method can be used by financial institutions for real time pricing of customized options and investment contracts with guarantees, and valuation under uncertainty of new ventures and the related growth financing instruments. We compare the proposed (hybrid) method with the simple use of ANN for very hard option pricing problems and we demonstrate the method's superiority. The combination of accurate option pricing and the resulting efficiency are instrumental for the applications we discuss.

1. Introduction

The methodology of contingent claims pricing has become one of the most (if not the most) significant contributions of financial economics, and allows valuation under uncertainty of standardized or customized financial derivatives. Main initial contributions to the literature were Black and Scholes (1973) and Merton (1973a). Since the original contributions, a vast number of models have appeared (see for example Briys et al, 1998, for some of the involved models and areas of application). The methodology of contingent claims pricing has also been extended to investment decision-making under uncertainty and the pricing of real (investment) options. A seminal contribution here is McDonald and Siegel (1986) and good review of an ever expanding literature can be found in Trigeorgis (1996), and Dixit and Pindyck (1994).

In all the above areas of applications only in the rare cases analytic solutions exist. In most cases (often computationally intensive) numerical methods must be used (see for example, Rogers and Talay, 1997, and Clewlow and Strickland, 1998). Recently, (nonparametric) Artificial Neural Networks (ANN) methods have been added to the toolkit of researchers and practitioners for valuation of derivatives (i.e., Hutchinson and Poggio, 1994; see also the review in Lajbcygier, 1999), valuation of real options (Taudes, et. al., 1998), and prediction of financial assets´ returns (i.e., Desai and Bharati, 1998). In this paper we propose the combined use of parametric and nonparametric (ANN) methods in order to price (customized) financial and real options efficiently, so that financial institutions and other intermediaries can implement them for purposes of real time decision making and valuation.

2. Function Approximation by Feedforward Neural Networks

Artificial neural networks (ANN) are massively parallel, highly connected structures consisting of a number of simple, nonlinear processing elements, called neurons; because of their massively parallel structure, can perform computations at very high rate if implemented on a dedicated hardware; because of their adaptive nature, they can learn the characteristics of input signals and adapt to changes in data; because of their nonlinear nature they can perform functional approximation which are beyond optimal linear techniques. Multilayer perceptions (MLP) are feedforward neural network with one or more layers of neurons, called hidden layers, between the output layer and the network's input. A three-layer feedforward network that is commonly used for approximation problems is shown in Fig. 1. It consists of an input layer and an output layer, corresponding to model input and output variables x and y, respectively, as well as a hidden layer.

Given the input feature vector x, the output can be computed by,

$$y(x) = v_0 + \sum_{i=1}^{H} v_i z_j$$

where z_i is the output of the ith hidden neuron computed as,

$$z_i = f(\psi_i)$$
$$\psi_i = w_{io} + \sum_{j=1}^{N} w_{ij} x_j$$

and f is the activation function. The two most widely activation functions are, Logistic: $$f(\psi) = \frac{1}{1+e^{-\psi}}$$

Tansigmoid: $$f(\psi) = \frac{1-e^{-\psi}}{1+e^{-\psi}}$$

The logistic activation function bounds *f,* in the range (0,1) while the tansigmoid function bounds *f* in the range (-1,1).

In the model development stage, samples of data (*x,y*) called training data, are generated from simulation or measurement. The neural network is then train by adjusting the weights w_{ij} and v_i, such that the neural network predicted output "best" matches that of training data, the target output. This is done by minimizing with respect to w_{ij} and v_i´s some norm of the error function between the predicted output of the neural network and the targeted outputs.

MLP were generally not popular due to the lack of effective learning algorithms, but this changed since the development of backpropagation (BP) learning algorithm proposed by Rumelhart, Hinton and Williams (1986). Although the BP algorithm is a simple learning algorithm for training MLP, it can exhibit oscillatory behavior or can even diverge; the algorithm is the steepest descent algorithm with predefined stepsize, the learning rate, λ.

MLP have the very important property (independent of the BP algorithm), namely that the computation of the gradient vector of its output can be accomplished by performing a forward analysis, with inputs being the training set, and a backward analysis, with inputs in the output layer being the errors obtained by the forward analysis. Hence we can replace the BP training algorithm by any of the powerful gradient optimization algorithms available in the literature, such as the conjugate gradient algorithms; Fletcher and Reeves (1964), Polak and Ribiere (see the book by Polak, 1971), Charalambous (1992), the Quasi-Newton algorithms; Fletcher (1970), Huang (1970).

In the model testing stage, a new set of input-output samples, called testing data, is used to test the accuracy of the neural model. The ability of neural models to give accurate y when presented with input values x never seen during training is called the generalization ability. A trained and tested neural network model can then be used online during option pricing evaluation providing fast model evaluation replacing the original simulators.

The usefulness of the above design approach is based on the following universal approximation theorem (see Cybenko, 1989).

Let f(.) be a nonconstant, bounded, and monotone-increasing continuous function. Let l_N denote the N-dimensional unit hypercube $[0,1]^N$. The space of continuous functions on l_N is denoted by $C(l_N)$. Then, given any function $g \in C(l_p)$ and $\varepsilon > 0$, *there exist an integer H and sets of real constants,*

$$w_{ij},\ v_i \quad i = 1,2,.....,H, \quad j = 0,1,...,N,$$

such that we may define

$$y(x) = \sum_{i=1}^{H} v_i f(\sum_{j=1}^{N} w_{ij} x_j + w_{io})$$

as an approximate realization of the function $g(.)$ *; that is,*

$$|y(x) - g(x)| < \varepsilon$$

for all vectors x that lie in the input space.

In effect, the theorem states that a single hidden layer is sufficient for a multilayer feedforward network to compute a uniform ε approximation to a given training set. However, the theorem does not say that a single hidden layer is optimum in the sense of learning time, or more importantly generalization. Moreover, the theorem does not give indications about the required number of hidden neurons necessary in order to achieve the desired degree of accuracy.

Several fundamental questions must be addressed in order to make practical use of the feedforward neural network models as approximations. A first question is the following: how many samples are needed to achieve a given degree of accuracy? It is well known that the answer depends on the dimensionality N of the data space and on the degree of smoothness of the function to be approximated. It has been shown that the number of samples grows exponentially with the ratio between the dimensionality N and the degree of smoothness.

An innovative approach, the hybrid approach, has been proposed by Watson and Gupta (1996) to reduce the training data needed and to improve the generalization capabilities of an ANN. In this model the target value for a given input vector *x*, is the difference in the response between that of the coarse model and that of the fine model. This leads into a simpler input-output relationship, and hence, reducing the number of fine model simulations.

3. Contingent Claims Pricing Problems

The option pricing literature was established with the seminal contribution by Black and Scholes (1973), and Merton (1973a). Subsequent work established the so-called risk-neutral valuation methodology and allows us to price claims dependent on several stochastic state-variables. Seminal contributions were Constantinides (1978), Harrison and Kreps (1979), Harrison and Pliska (1981), and Cox, Ingersoll, and Ross (1985). For review see Malliaris and Brock (1982), and Karatzas and Shreve (1991). Following the extended (with a dividend yield) Black and Scholes assumptions a traded asset can follow a stochastic process of the form

$$\frac{dS}{S} = \mu dt + \sigma_S dz_S$$

in the real probability measure, and

$$\frac{dS}{S} = (r - \delta_S)dt + \sigma_S dz_S$$

in the risk-neutral probability measure. The asset has a continuous expected growth rate of μ and a standard deviation of σ_S, it pays a continuous dividend yield of δ_S, with r being the continuous riskless rate of interest, and dz being the increment to the standard Wiener process. For simplicity we will present the valuation method for two (potentially foreign assets) and the method is easily extendible to more. We assume that there exist two assets in two foreign countries a and b, and we follow the notation that asset AR is in country a with exchange rate ER from our perspective, and asset AW is in country b with exchange rate EW. All state-variables follow geometric Brownian motion processes. The foreign currencies follow (see Garman and Kohlhagen, 1983, and Grabbe, 1983) the stochastic processes

$$\frac{dER}{ER} = (r - r_a)dt + \sigma_{ER} dz_{ER},$$

and

$$\frac{dEW}{EW} = (r - r_b)dt + \sigma_{EW} dz_{EW},$$

where r is the local, and r_a, r_b are the foreign (constant) riskless rates of interest, standard deviations σ_{ER} and σ_{EW}, and covariance $\sigma_{ER,EW}$. The two foreign assets from the perspective of the option holder but before they are translated to the option holder's currency (see Reiner, 1992, and Kat and Roozen, 1994) follow the risk-neutral process

$$\frac{dAR}{AR} = (r_a - \delta_{AR} - \sigma_{AR,ER})dt + \sigma_{AR} dz_{AR},$$

and

$$\frac{dAW}{AW} = (r_b - \delta_{AW} - \sigma_{AW,EW})dt + \sigma_{AW} dz_{AW}.$$

We see that the risk-neutral drifts include not only the (local) dividend yields δ_{AR} and δ_{AW}, but also the instantaneous covariance between the exchange rate and the cash flow (see *Siegel's paradox* in Hull, 1997, pp. 298-301). It is known that the partial differential equation (PDE) for the claim V dependent on AR, ER, AW, and EW is

$$\begin{aligned}
&0.5\sigma_{AR}^2 AR^2 \frac{\partial^2 V}{\partial AR^2} + 0.5\sigma_{AW}^2 AW^2 \frac{\partial^2 V}{\partial AW^2} + 0.5\sigma_{ER}^2 ER^2 \frac{\partial^2 V}{\partial ER^2} + 0.5\sigma_{EW}^2 EW^2 \frac{\partial^2 V}{\partial EW^2} \\
&+ \sigma_{AR,ER}^2 ARER \frac{\partial^2 V}{\partial AR \partial ER} + \sigma_{AW,EW}^2 AWEW \frac{\partial^2 V}{\partial AW \partial EW} + \sigma_{AR,EW}^2 AREW \frac{\partial^2 V}{\partial AR \partial EW} \\
&+ \sigma_{ER,AW}^2 ERAW \frac{\partial^2 V}{\partial ER \partial AW} + \sigma_{AR,AW}^2 ARAW \frac{\partial^2 V}{\partial AR \partial AW} + \sigma_{ER,EW}^2 EREW \frac{\partial^2 V}{\partial ER \partial EW} \\
&+ (r_a - \delta_{AR} - \sigma_{AR,ER}) AR \frac{\partial V}{\partial AR} + (r_b - \delta_{AW} - \sigma_{AW,EW}) AW \frac{\partial V}{\partial AW} \\
&+ (r - r_a) ER \frac{\partial V}{\partial ER} + (r - r_b) EW \frac{\partial V}{\partial EW} = rV
\end{aligned}$$

In the presence of more foreign assets this PDE will be extended with two variables for each new asset in a new foreign country (all other combinations are special cases; see also Martzoukos, 1997). For simplicity we present the PDE for two foreign assets only and in the presence of more assets the PDE is extended similarly. Separation of the assets on the underlying state-variables as above allows solution of complex problems involving local or foreign stocks, foreign currencies, or assets protected from exchange rate moves through contractual provisions. To simplify the exposition we will assume that the underlying assets are simply the foreign assets without any exchange rate protection, and with the use of standard Ito calculus tools we reduce the PDE from one of four state-variables to one of two asset prices, the R and W (or S_1, S_2, etc.). Using the multi-dimensional form of Ito´s lemma it can be easily shown that from the perspective of the option holder and under risk-neutrality

$$\frac{dR}{R} = (r - \delta_{AR})dt + \sigma_R dz_R ,$$

and

$$\frac{dW}{W} = (r - \delta_{AW})dt + \sigma_W dz_W .$$

The two-dimensional PDE finally is

$$0.5\sigma_R^2 R^2 \frac{\partial^2 V}{\partial R^2} + 0.5\sigma_W^2 W^2 \frac{\partial^2 V}{\partial W^2} + \sigma_{R,W}^2 RW \frac{\partial^2 V}{\partial R \partial W}$$
$$+ (r - \delta_{AR})R\frac{\partial V}{\partial R} + (r - \delta_{AW})W\frac{\partial V}{\partial W} = rV$$

with

$$\sigma_R^2 = \sigma_{AR}^2 + \sigma_{ER}^2 + 2\sigma_{AR,ER},$$

$$\sigma_W^2 = \sigma_{AW}^2 + \sigma_{EW}^2 + 2\sigma_{AW,EW},$$

$$\sigma_{R,W} = \sigma_{AR,AW} + \sigma_{AR,EW} + \sigma_{ER,AW} + \sigma_{ER,EW}.$$

If more than two underlying assets exist, the above PDE is expanded accordingly. Analytic solution of such PDEs is rather unlikely although possible for very special cases. In the most general case numerical solutions will be needed, either direct (implicit or explicit numerical discretization methods), or indirect by use of option pricing techniques that simulate the underlying stochastic processes like Monte-Carlo simulation or lattice-based methods, etc. The option-pricing problem is finally solved by taking the appropriate boundary conditions into account. Such conditions relate to the specification of the contingent claim as a call or put option, European or American (allowing for early exercise), option on the maximum or the average of the underlying assets, and other assumptions, for example contractual specification of path-dependency, etc. The *first* example that we use in this paper is the European option on the average of three assets, and for pricing we implement a three-dimensional lattice.

Real Investment Options: The Case of Higher-Dimensional Operational Flexibility with (Path-Dependency Inducing) Switching Costs.

In the real option models, investment decision-making and valuation of a claim (or *real option*) on risky ventures is contingent on stochastic (real) asset(s), following often geometric Brownian motion process(es) with constant drift(s) and instantaneous variance(s). The methods of valuation have been established in the literature that applies stochastic calculus to valuation of real options, with a seminal publication that of McDonald and Siegel (1986), who consider the real option to wait-to-invest (in the context of complete irreversibility and a single investment decision). Applications of this standard model have appeared in Crousillat and Martzoukos (1991). Partial irreversibility models that allow reversible (and costly) investment and disinvestment decisions were first introduced in the literature by Brennan and Schwartz (1985). In general the real options literature has demonstrated that the classic Net Present Value

(NPV) rule fails under uncertainty and irreversibility-inducing sunk costs. Review of this literature can be found in Pindyck (1991), Dixit (1992), Dixit and Pindyck (1994), and Trigeorgis (1996). For more recent contributions to the literature, see for example Trigeorgis (1993), Tannous (1996), Alvarez (1999), Childs and Triantis (1999), and Dangl (1999).

In general, for real options pricing we draw on Constantinides (1978) or Cox, Ingersoll, and Ross (1985), and a continuous time capital asset pricing model (see Merton, 1973b, or Breeden, 1979) is assumed to hold. The underlying asset is not necessarily a traded one. The difference between the required rate of return on the underlying asset and its growth rate is denoted by δ, and is an opportunity cost of deferring investment in the capital intensive project (see McDonald and Siegel, 1984); for a convenience yield interpretation, see Brennan and Schwartz (1985), and Brennan (1991).

We are interested in the very hard problems of partial irreversibility like the ones introduced by Brennan and Schwartz (1985) in a model for natural resource investment decisions, because the costs of switching among different modes of operation induce path-dependency within the *zone of inaction*. Specifically, this zone is a range of asset prices where investment (or switching) decisions are path-dependent. In this zone, if we have not invested, we remain so; likewise, if we have invested, we remain invested. This path dependency makes the valuation of such investment options difficult. Soon thereafter, Dixit (1989a, 1989b, 1989c) generalized their results. Dixit (1989a, c) used the model to show the path-dependency effects in a simplified two-sector economy with costly capital mobility, and Dixit (1989b) used the model to show the effects in entering in, or exiting from, a foreign market when the exchange rate is risky and follows geometric Brownian motion. This model considers the more general case of industry equilibrium when the company has the (sequential) option to purchase or abandon many existing producers. Extensions of these models have also appeared in corporate finance (Mauer and Triantis, 1994), and under foreign exchange as the sole factor of risk (Mello, Parsons, and Triantis, 1995, and Bell, 1995). The *second* example we use in this paper is the valuation of a flexible operation that offers the option to operate on the best of two (stochastic) assets, with costly switching between the operating and the idle states.

4. Two Examples

In this work, two examples will be considered to show the usefulness of ANN to option pricing. *First* we assume an investor with a position promising to pay a return linked to the average of three assets (local or foreign stocks or stock indexes). If in addition the investor is offered protection so that the value of the initial investment cannot fall below a certain level, the value of this protection is equivalent to the value of a put option on the average of the three assets. Such options are contractual guarantees embedded in the investment products the financial institutions offer to investors. In the absence of path-dependency European and American options can be priced using lattice-based methods like the Boyle, Evnine, and Gibbs (1989) that we implement (see also Kamrad and Rithcken, 1991, and Ekvall, 1996). In order to have accurate results, it is known that we need to use a relatively dense (with many steps) lattice, thus increasing the computational intensity of the multi-dimensional problem. The price we derive (the *fine* model) is the average of those provided by a 40-step and a 41-step lattice. If we use the *coarse* model alone (for efficiency purposes), the price is derived from a 10-step lattice. The difference in computational intensity between the two is by a factor of 357 (option node evaluations). Since the accuracy standard is the computationally intensive *fine* model, we will also use ANN to *efficiently* price these financial options. We allow six variables to take five different values each, thus creating a *training* set of 7776 for each of the *fine* and the *coarse* option values, and each of the input variables. The input variables are the three underlying assets S_1, S_2 and S_3, each with values equally spaced between 85 and 115, the riskless rate of interest r with values between 0.02 and 0.06, all correlations of the assets′ continuous rates of return $\rho_{1,2}$, $\rho_{1,3}$ and $\rho_{2,3}$ (equal to each other for simplicity of exposition) between –0.20 and 0.20, and all standard deviations of the assets′ continuous rates of return σ_1, σ_2 and σ_3 (again equal to each other) between 0.10 and 0.50. The exercise price is fixed at $X = 100$, the two dividend yields δ_1 and δ_2 are fixed at 0.03, and the time to maturity fixed at $T = 3.00$. We also create a (out-of-the-sample) *test* set of 200 by drawing randomly and independently for the six input variables from uniform distributions in the range of the values used in the *training* set.

Our *second* application is in the context of real options pricing with operating flexibility and switching costs. Such types of investment models can capture in general the effects of multiple uncertainties (several underlying assets, local or in foreign countries); the flexibility to choose among several products (output) or raw material (input); and the flexibility to operate or not according to the profitability of the operation (switching decisions equivalent to the early exercise feature of an American-type option). The difficulty of pricing under uncertainty this American-type contingent claim and of determining the optimal state is because of the path-dependency inducing switching costs.

Our specific example involves pricing a risky venture with the operating flexibility to choose the production of the best of *two* outputs (with stochastic values) by paying a fixed production cost, and to switch between the active (operating) and the idle mode after paying fixed switching costs. In the most general case, both underlying assets (the outputs) can be multiplicative functions of two uncertainties, an asset price and an exchange rate. The continuous one-dimensional case would be similar to Dixit (1989a) in a local (one-country) setting, or Martzoukos (1998) in a multinational (three-country) framework. The solution of such problems is very hard (for an early reference see Kulatilaka, 1988). The method we implement is equivalent to solving a multi-stage optimization problem by implementing a recursive forward-backward looking algorithm of exhaustive search. Alternatively but similarly, a procedure like in Hull and White (1993) can be implemented (by using a 2-D lattice instead of the 1-D discussed therein). The *fine* model allows optimal decisions (switching) at time zero, at the maturity of the option, and twice in-between, thus the optimization problem with exhaustive search is similar to a 3-stage stochastic programming model. Note that the decision at maturity is also path-dependent on the previous decisions, so the problem is practically a 4-stage one, but at the 4^{th} stage since we have reached option maturity the optimization problem conditional on the path is not a stochastic one. We implement a 24-step 2-D lattice for the *fine* model and an 8-step 2-D lattice for the *coarse* model. The difference in computational intensity between the two is by a factor of 1567 (option node evaluations). Since the accuracy standard is the extremely computationally intensive *fine* model, we also use ANN to enhance the *efficiency* of calculations. We allow five variables to take five different values each, thus creating a *training* set of 3125 for each of the *fine* and the *coarse* option values, and each of the input variables.

The input variables are the two underlying assets (product values) S_1 and S_2, each with values equally spaced between 85 and 115, the standard deviations of the assets′ continuous rates of return σ_1 and σ_2 (equal to each other) between 0.10 and 0.50, the switching cost to get to active mode (from idle) between 5.00 and 25.00, and the switching cost to get to idle mode (from active) similarly between 5.00 and 25.00. The exercise price is fixed at $X = 100$, the riskless rate of interest r is fixed at 0.05, the two dividend yields δ_1 and δ_2 are fixed at 0.05, the correlation of the assets′ continuous rates of return $\rho_{1,2}$ is fixed at 0.35, and the time to maturity fixed at $T = 3.00$. We also create a *test* set of 200 with random drawing exactly like in the first example.

For both examples the feedforward neural network structure shown in Fig. 1 with one hidden layer will be used.

The performance of the two neuromodels shown in Figs. 2 and 3 will be investigated; the simple neuromodel shown in Fig. 2 tries to capture the functional relationship between the option price (fine model) and the input variables, while the neuromodel shown in Fig.3 tries to capture the functional relationship between the deviation of the option price (fine model) from that given by the coarse model and the input variables. The neuromodel shown in Fig. 2 will be called simple ANN model. The neuromodel shown in Fig. 3 is based on the hybrid model of Watson and Gupta and will be called Hybrid Numerical Option Pricing and ANN (NOP-ANN) model.

The Quasi-Newton, BFGS algorithm (see for example Fletcher, 1980) in conjunction with Charalambous (1992) line search is used to efficiently train the ANN′s as a powerful alternative to popular backpropagation algorithm.

Tables 1 and 2 show the results obtained for problem 1 and 2, respectively. For the hybrid model we considered different values for the number H of neurons in hidden layer, ranging from 1 to 20, while for the simple model we considered only the case where the number of neurons in the hidden layer is 20. Three measures are used to compare the results:

mse (e): mean square error

mae (e): mean absolute error

max (e): maximum absolute error

For both neuromodels, error is defined as the difference between the option price obtained by the neuromodels and that of a fine model. The elements in brackets correspond to the results obtained on the training data, and the rest of the elements correspond to the results obtained on the testing data. Examining the tables we can conclude the following:

- The simple model, even with 20 neurons in the hidden layer, did not give much better results that the ones obtained by the coarse model.
- The results obtained by the hybrid model with only one neuron in the hidden layer are slightly better than those obtained by the coarse model; this should be expected, because in a way in the hybrid approach, we are taking the coarse model as our based model.
- The results obtained by the hybrid model with 20 neurons in the hidden layer are superior to those obtained by the coarse model.

Figures 4 and 5 show plots of the error for the 200 testing sample points, for both the hybrid model with $H = 20$ and the coarse model.

Summary and conclusions

In this paper we propose a method for fast and efficient (practically real time) pricing of contingent claims through the use of ANN. We demonstrate results for financial derivatives and real (investment) options. Specifically we price customized investment vehicles with embedded guarantees offered by financial institutions to individuals, and risky ventures with (costly) switching of modes of operation due to the strategic flexibility of dynamic management of the entrepreneur. Other contracts can be similarly priced, like complex warrants, venture capital financing vehicles (i.e., convertibles), etc. The results demonstrate tremendous advantages in computational efficiency through the combined use of option pricing models and ANN. The use of such intensive option pricing methods is thus made available for real time applications where institutions´ financial officers investigate alternative contractual terms and/or discuss them with clients. Some of these methods can be made available to current or prospective clients on the internet.

References

Alvarez, L.H.R., "Optimal Exit and Valuation under Demand Uncertainty: A Real Options Approach." *European Journal of Operational Research* 114, 320-329, (1999).

Bell, K.G., Volatile Exchange Rates and the Multinational Firm: Entry, Exit, and Capacity Options. In *Real Options in Capital Investment: Models, Strategies, and Applications*, ed. Trigeorgis, Westport, Connecticut: Praeger, 1995.

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, 637–659, (1973).

Boyle, P.P., J. Evnine, and S. Gibbs, "Numerical Evaluation of Multivariate Contingent Claims." *Review of Financial Studies* 2, 241-250, (1989).

Breeden, D.T., "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." *Journal of Financial Economics* 7, 265–296, (1979).

Brennan, M.J., The Price of Convenience and the Valuation of Commodity Contingent Claims. In *Stochastic Models and Option Values*, eds. D. Lund and B. Öksendal, New York: North-Holland, 1991.

________ and E. Schwartz, "Evaluating Natural Resource Investments." *Journal of Business* 58, 135–157, (1985).

Briys, E., M. Bellalah, H. M. Mai, and F.D. Varenne, *Options, Futures and Exotic Derivatives*, John Wiley & Sons, 1998.

Charalambous, C., "A conjugate Gradient Algorithm for Efficient Training of Artificial Neural Networks", IEE Proceedings – G, 139, 301-310, (1992).

Childs, P.D., and A.J. Triantis, "Dynamic R&D Investment Policies." *Management Science* 45, 1359-1377, (1999).

Clewlow, L., and C. Strickland, *Implementing Derivatives Models*, John Wiley & Sons, 1998.

Constantinides, G., "Market Risk Adjustment in Project Valuation." *Journal of Finance* 33, 603–616, (1978).

Cox, J., J. Ingersoll, and S. Ross, "An Intertemporal General Equilibrium Model of Asset Prices." *Econometrica* 53, 363–384, (1985).

Crousillat, E. and S.H. Martzoukos, "Decision Making under Uncertainty - An Option Valuation Approach to Power Planning." IENED # 39, The World Bank, (1991).

Cybenko, G., "Approximation by Superpositions of a Sigmoidal Function", *Mathematics of Control, Signals and Systems* 2, 303-314, (1989).

Dangl, T., "Investment and Capacity Choice under Uncertain Demand." *European Journal of Operational Research* 117, 415-428, (1999).

Desai, V.S., and R. Bharati, "The Efficacy of Neural Networks in Predicting Returns on Stock and Bond Indices." *Decision Sciences Journal* 29, (1998).

Dixit, A., "Entry and Exit Decisions under Uncertainty." *Journal of Political Economy* 97, 620–638, (1989a).

_____, "Hysteresis, Import Penetration, and Exchange Rate Pass-Through." *Quarterly Journal of Economics* 104, 205–228, (1989b).

_____, "Intersectoral Capital Allocation Under Price Uncertainty." *Journal of International Economics* 26, 309–325, (1989c).

_____, "Investment and Hysteresis." *Journal of Economic Perspectives* 6, 107–132, (1992).

_____ and R. Pindyck, *Investment under Uncertainty*. Princeton NJ: Princeton University Press, 1994.

Ekvall, N., "A Lattice Approach for Pricing Multivariate Contingent Claims." *European Journal of Operational Research* 91, 214-228, (1996).

Fletcher R., "A New Approach to Variable Metric Algorithms", *Computer Journal* 13, 317-322, (1970).

Fletcher, R., "Practical Methods of Optimization", John Wiley & Sons, 1980.

Fletcher, R., and Reeves, C.M., "Function Minimization by Conjugate Gradients" *Computer Journal* 7, 149-154, (1964).

Garman, M. and S.W. Kohlhagen, "Foreign Currency Option Values." *Journal of International Money and Finance* 2, 231-237, (1983).

Grabbe, O.J., "The Pricing of Call and Put Options on Foreign Exchange." *Journal of International Money and Finance* 2, 239-253, (1983).

Harrison, J.M. and D. Kreps, "Martingales and Arbitrage in Multiperiod Security Markets." *Journal of Economic Theory* 20, 381–408, (1979).

Harrison, J.M. and S. Pliska, "Martingales and Stochastic Integrals in the Theory of Continuous Trading." *Stochastic Processes and Their Applications* 11, 313–316, (1981).

Huang, H.Y., "Unified Approach to Quadratically Convergent Algorithms for Function Minimization." *Journal of Optimization Theory and Applications* 5, 405–423, (1970).

Hull, J., *Options, Futures, and other Derivatives*, Upper Saddle River, New Jersey: Prentice Hall, 1997.

Hull, J., and A. White, "Efficient Procedures for Valuing European and American Path-Dependent Options." *Journal of Derivatives* 1, 21-31, (1993).

Hutchinson, J., A. Lo, and T. Poggio, "A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks." *Journal of Finance* 49, 851-889, (1994).

Kamrad, B., and P. Ritchken, "Multinomial Approximating Models for Options with k State-variables." *Management Science* 37, 1640-1652, (1991).

Karatzas, I. and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Berlin, Germany: Springer-Verlag, 1991.

Kat, H.M. and H.N.M. Roozen, "Pricing and Hedging International Equity Derivatives." *Journal of Derivatives* 2, 7–19, (1994).

Kulatilaka, N., "Valuing the Flexibility of Flexible manufacturing Systems." *IEEE Transactions on Engineering Management* 35, 250-257, (1988).

Lajbcygier, P., "Literature Review: The Problem with Modern Parametric Option Pricing." *Journal of Computational Intelligence in Finance*, 6-23 (September/October 1999).

Malliaris, A.G. and W. Brock, *Stochastic Methods in Economics and Finance*, New-York: North-Holland, 1982.

Martzoukos, S.H., Multivariate Contingent Claims with Stochastic Exercise Price and Multiplicative Uncertainties. George Washington University, SBPM working paper 96–42 (revision 1997), forthcoming at the *Journal of Multinational Financial Management*.

Martzoukos, S.H., "Hysteresis Models of Investment with Multiple Uncertainties and Exchange Rate Risk." Working paper (1998), forthcoming at the *Review of Quantitative Finance and Accounting*.

Mauer, D. and A. Triantis, "Interactions of Corporate Financing and Investment Decisions: A Dynamic Framework." *Journal of Finance* 49, 1253–1277, (1994).

McDonald, R. and D. Siegel, "Option Pricing when the Underlying Asset Earns a Below-Equilibrium Rate of Return: A Note." *Journal of Finance* 39, 261–265, (1984).

__________, "The Value of Waiting to Invest." *Quarterly Journal of Economics* 101, 707–727, (1986).

Mello, A.S., J.E. Parsons, and A.J. Triantis, "An Integrated Model of Multinational Flexibility and Financial Hedging." *Journal of International Economics* 39, 27–51, (1995).

Merton, R., “The Theory of Rational Option Pricing.” *Bell Journal of Economics and Management Science* 4, 141–183, (1973a).

______, “An Intertemporal Capital Asset Pricing Model.” *Econometrica*, 867–887, (1973b).

Pindyck, R., “Irreversibility, Uncertainty and Investment.” *Journal of Economic Literature* 29, 1110–1152, (1991).

Polak, E., *Computational Methods in Optimization: A Unified Approach*. New York: Academic Press, 1971.

Reiner, E., “Quanto Mechanics.” *RISK* March, 59–63, (1992).

Rogers, L.C.G., and D. Talay, editors, *Numerical Methods in Finance*. Cambridge: Cambridge University Press, 1997.

Rumelhart, D., Hinton, G., and Williams, G, “Learning Internal representations by error propagation”, in Rumelhart, D., and McCleland, J. (Ed.): “Parallel Distributed Processing”, vol. 1 (MJT Press, 1986).

Tannous, G.F., “Capital Budgeting for Volume Flexible Investment.” *Decision Sciences Journal* 27, (1996).

Taudes, A., M. Natter, and M. Trcka, "Real Option valuation with Neural Networks." *International Journal of Intelligent Systems in Accounting, Finance & Management* 7, 43-52 (1998).

Trigeorgis, L., "The Nature of Option Interactions and the Valuation of Investments with Multiple Real Options." *Journal of Financial and Quantitative Analysis* 28, 1–20 (1993).

__________, *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. Cambridge: The MIT Press, 1996.

Watson, P., and Gupta, K.C., "EM-ANN models for microstrip vias and Interconnect in Multilayer Circuits", *IEEE Trans. microwave Theory and Techniques* 44, 2495-2503 (1996).

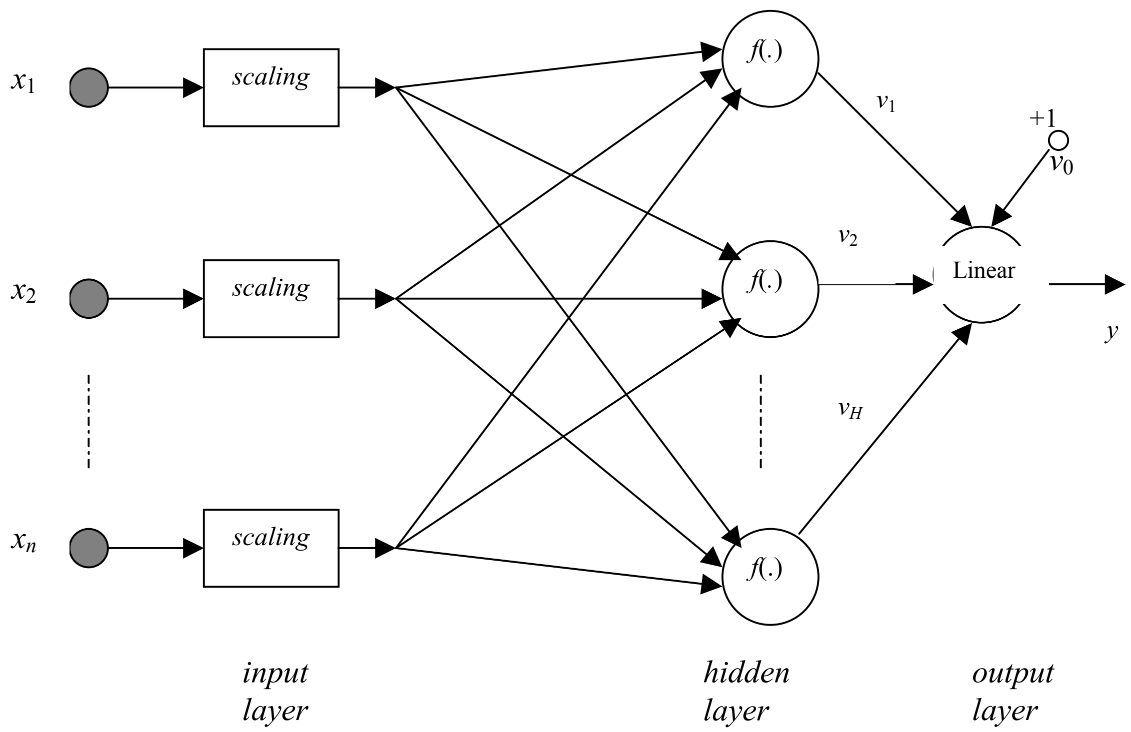

Figure 1: Feedforward ANN model with a single hidden layer

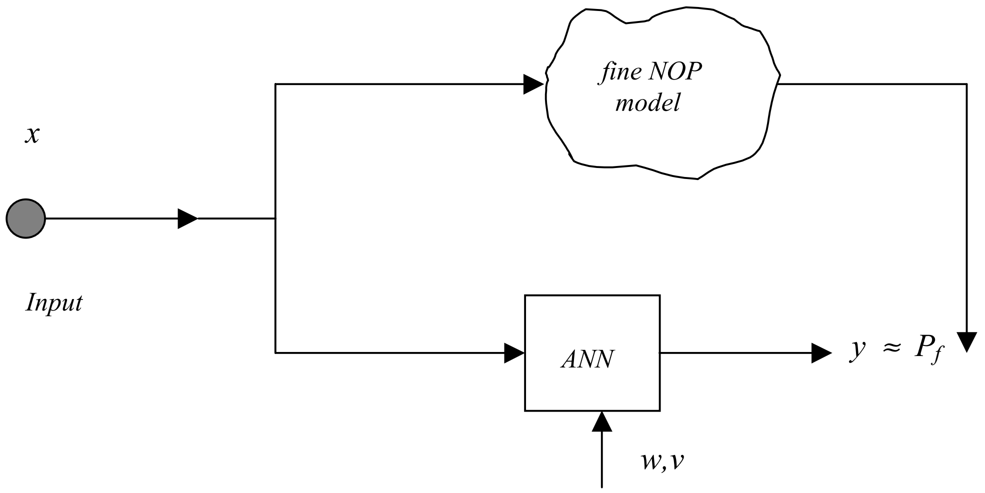

Figure 2: Simple ANN model

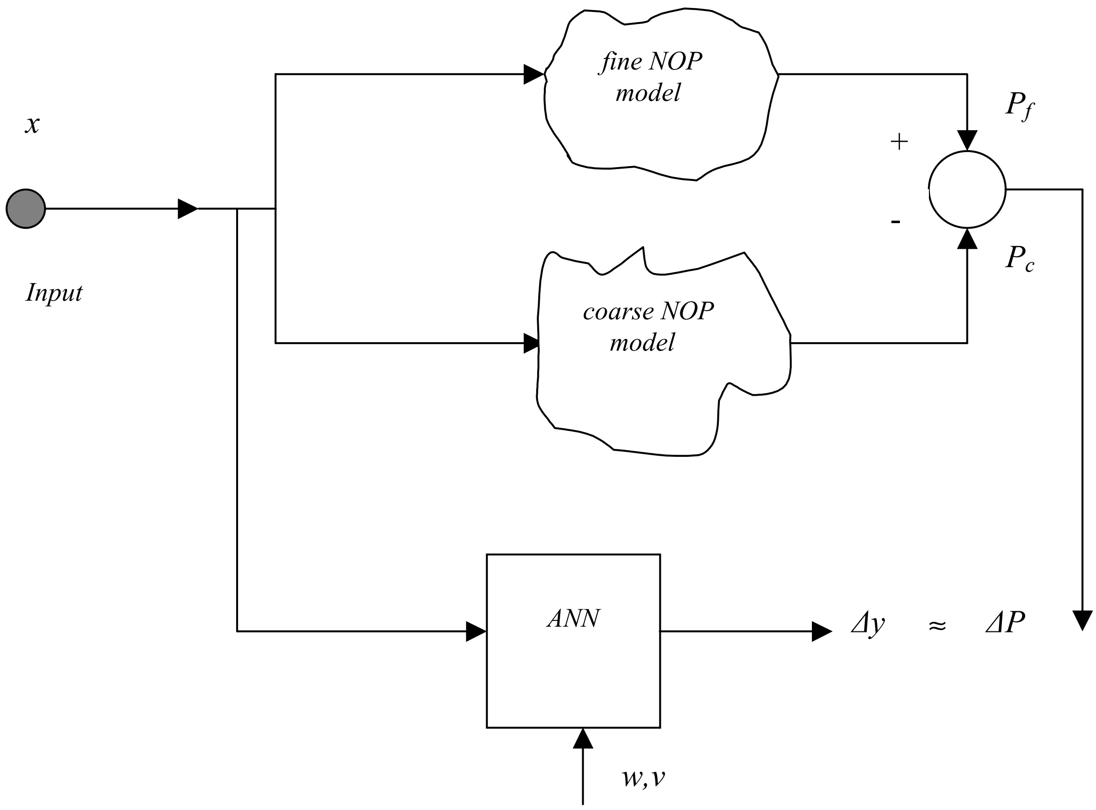

Figure 3: Hybrid Numerical Option Pricing and ANN model

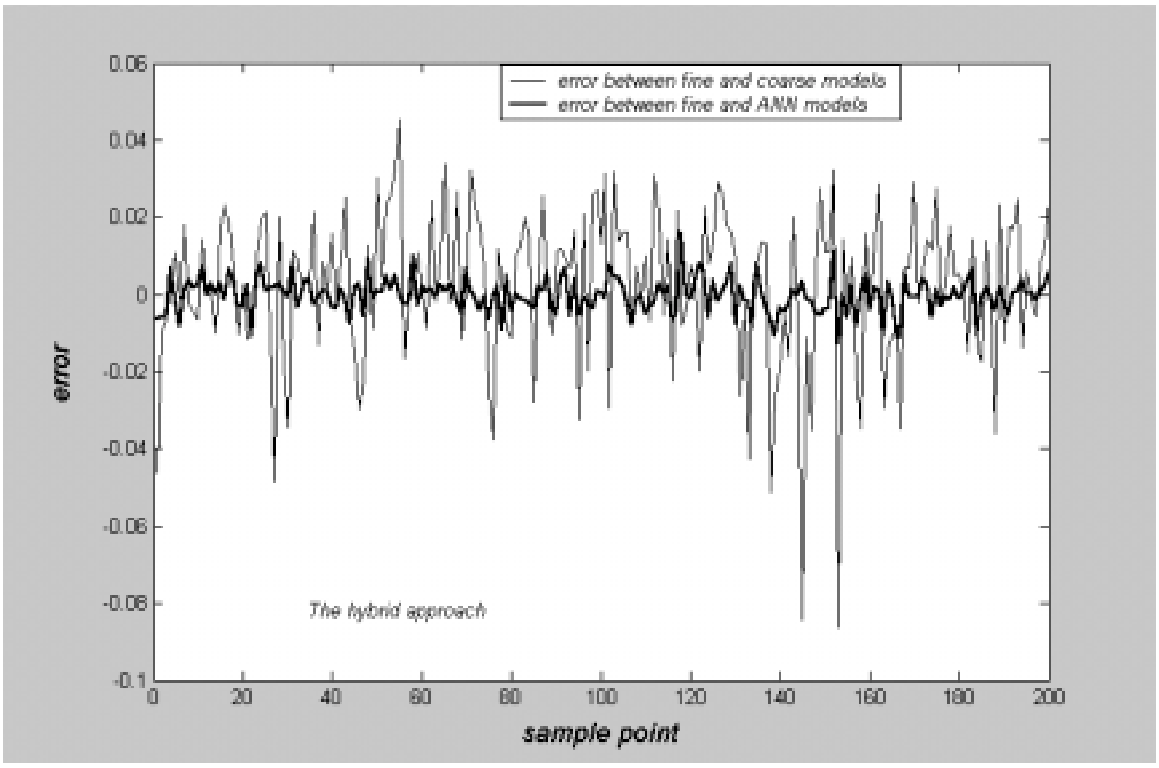

Figure 4: **Plot of the error for the testing sample points, for both the hybrid model with H=20 and the coarse model for example 1**

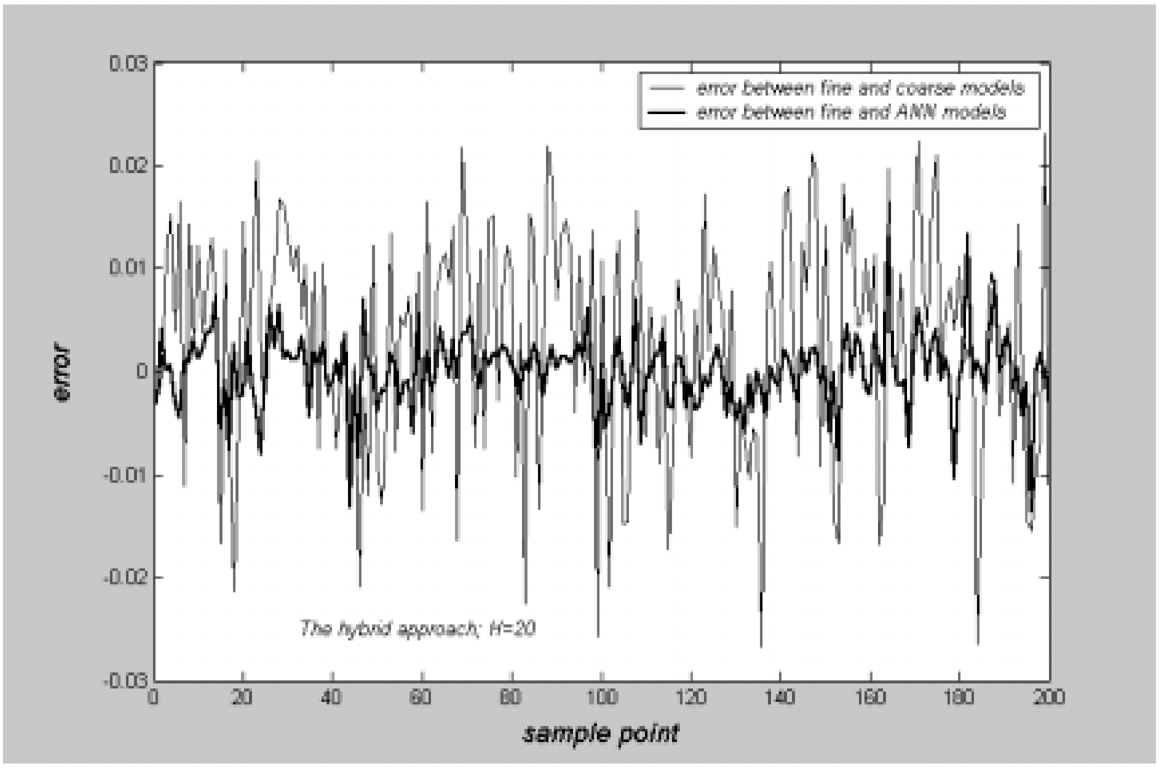

Figure 5: **Plot of the error for the testing sample points, for both the hybrid model with H=20 and the coarse model for example 2**

Error Measure	***Hybrid NOP-ANN Model***					***Simple ANN Model***	***Coarse NOP Model***
	H					*H*	
	1	2	5	10	20	20	
$mse(e)$x10^4	(3.2) 2.5	(1.1) 0.9	(0.82) 0.74	(0.39) 0.60	(0.23) 0.21	(2.7) 1.8	(6.0) 4.1
$mae(e)$x10^2	(1.4) 1.3	(0.85) 0.79	(0.75) 0.72	(0.47) 0.59	(0.36) 0.35	(1.2) 1.1	(1.8) 1.6
$max\ \lvert e\rvert$x10^2	(8.1) 6.0	(8.3) 3.4	(7.4) 2.9	(5.0) 2.1	(2.7) 1.6	(2.3) 4.4	(11.4) 8.6

Table 1: ***Comparison between the hybrid NOP-ANN, the simple ANN, and the coarse NOP model on example 1. The 1st and 2nd rows in each measure correspond to training and testing sets respectively.***

Error Measure	***Hybrid NOP-ANN Model***					***Simple ANN Model***	***Coarse NOP Model***
	H					*H*	
	1	2	5	10	20	20	
mse(*e*)x10^4	(1.1) 1.0	(0.79) 0.85	(0.47) 0.45	(0.24) 0.21	(0.16) 0.14	(1.5) 1.0	(1.3) 1.3
mae(*e*)x10^2	(0.84) 0.84	(0.72) 0.75	(0.55) 0.54	(0.38) 0.33	(0.30) 0.27	(0.97) 0.79	(0.90) 0.95
max |*e*|x10^2	(4.4) 3.2	(3.3) 2.3	(2.4) 2.2	(2.1) 1.9	(2.1) 1.4	(6.1) 2.8	(4.7) 2.7

Table 2: ***Comparison between the hybrid NOP-ANN, the simple ANN, and the coarse NOP model on example 2. The 1st and 2nd rows in each measure correspond to training and testing sets respectively.***

WORKING PAPER SERIES

HERMES Center of Excellence
on Computational Finance & Economics
University of Cyprus

Preprints of these working papers can be downloaded from www.hermes.ucy.ac.cy or obtained by sending an e-mail to hermes@ucy.ac.cy

Working Paper 01-01	*STOCHASTIC PROGRAMMING MODELS FOR ASSET AND LIABILITY MANAGEMENT* *Roy Kouwenberg and Stavros A. Zenios*
Working Paper 01-02	*THE VALUE OF INTEGRATIVE RISK MANAGEMENT FOR INSURANCE PRODUCTS WITH GUARANTEES* *Andrea Consiglio, Flavio Cocco and Stavros A. Zenios*
Working Paper 01-03	*EXTENDING CREDIT RISK (PRICING) MODELS FOR THE SIMULATION OF PORTFOLIOS OF INTEREST RATE AND CREDIT RISK SENSITIVE SECURITIES* *Norbert Jobst and Stavros A. Zenios*
Working Paper 01-04	*TRACKING CORPORATE BOND INDICES IN AN INTEGRATED MARKET AND CREDIT RISK ENVIRONMENT* *Norbert Jobst and Stavros A. Zenios*
Working Paper 01-05	*THE TAIL THAT WAGS THE DOG: INTEGRATING CREDIT RISK IN ASSET PORTFOLIOS* *Norbert Jobst and Stavros A. Zenios*
Working Paper 01-06	*PARALLELIZATION, OPTIMIZATION, AND PERFORMANCE ANALYSIS OF PORTFOLIO CHOICE MODELS* *Ahmed Abdelkhalek, Angelos Bilas, and Alexander Michaelides*
Working Paper 01-07	*REAL R&D OPTIONS WITH ENDOGENOUS AND EXOGENOUS LEARNING* *Spiros H. Martzoukos*
Working Paper 01-08	*PORTFOLIO CHOICE AND LIQUIDITY CONSTRAINTS* *Michael Haliassos and Alexander Michaelides*

(Also published as CEPR Discussion Paper No. 2822)

Working Paper 01-09	*PORTFOLIO CHOICE, LIQUIDITY CONSTRAINTS AND STOCK MARKET MEAN REVERSION* *Alexander Michaelides*

(Also published as CEPR Discussion Paper No. 2823)

Working Paper 01-10 *BORROWING CONSTRAINTS, PORTFOLIO CHOICE, AND PRECAUTIONARY MOTIVES*
Michael Haliassos and Christis Hassapis

Working Paper 01-11 *DEBT REVOLVERS FOR SELF CONTROL*
Carol C. Bertaut and Michael Haliassos

Working Paper 01-12 *EQUITY CULTURE AND HOUSEHOLD BEHAVIOR*
Michael Haliassos and Christis Hassapis

Working Paper 01-13 *SCENARIO OPTIMIZATION ASSET AND LIABILITY MODELING FOR ENDOWMENTS WITH GUARANTEES*
Andrea Consiglio, Flavio Cocco and Stavros A. Zenios
(Also published as WFIC Working Paper 00-41-B)

Working Paper 01-14 *RESOLVING A REAL OPTIONS PARADOX WITH INCOMPLETE INFORMATION: AFTER ALL, WHY LEARN?*
Spiros H. Martzoukos and Lenos Trigeorgis

Working Paper 01-15 *REAL OPTIONS WITH INCOMPLETE INFORMATION AND MULTI-DIMENSIONAL RANDOM CONTROLS*
Spiros H. Martzoukos

Working Paper 01-16 *HOUSEHOLD PORTFOLIOS: AN INTERNATIONAL COMPARISON*
Luigi Guiso, Michael Haliassos and Tullio Jappelli

Working Paper 01-17 *CALIBRATION AND COMPUTATION OF HOUSEHOLD PORTFOLIO MODELS*
Michael Haliassos and Alexander Michaelides

Working Paper 01-18 *ENTERPRISE-WIDE ASSET AND LIABILITY MANAGEMENT: ISSUES, INSTITUTIONS, AND MODELS*
Dan Rosen and Stavros A. Zenios

Working Paper 01-19 *ARTIFICIAL NEURAL NETWORKS FOR VALUATION OF FINANCIAL DERIVATIVES AND CUSTOMIZED OPTION EMBEDDED CONTRACTS*
Christakis Charalambous and Spiros H. Martzoukos

Working Paper 01-20 *REAL (INVESTMENT) OPTION GAMES WITH INCOMPLETE INFORMATION AND LEARNING SPILLOVERS*
Spiros H. Martzoukos and Eleftherios Zacharias